Identification and characterisation of a novel *Plasmodium falciparum* protein.

Louise Hinds

University College London,
University of London

PhD

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

I, Louise Hinds, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed: _

Abstract

The aim of this project was to identify and characterise a novel *Plasmodium falciparum* protein that plays a role in the invasion of host cells. Published genome, proteome and transcriptome data were used to select a novel protein for study. The data was filtered according to a set of criteria designed to identify uncharacterised proteins with features that would be expected of an invasion molecule, e.g. presence of a signal peptide and membrane anchor, as well as a late blood-stage expression profile. The gene product encoded by *PF08_0008* was ultimately selected for study.

Two regions of PF08_0008, from the N- and C-termini, were expressed in a recombinant form in *E. coli* and used to generate specific polyclonal antibodies in mice and rabbits. These antibodies have been used in a variety of experiments to characterise PF08_0008. Indirect immunofluorescence assays have shown that PF08_0008 is located at the parasite's apical end in both segmented schizonts and free merozoites. Immunoprecipitation and Western blot analyses have shown that PF08_0008 is proteolytically processed prior to schizont rupture. Fragments of PF08_0008 are detectable in culture supernatant, suggesting that shedding of this protein from the merozoite surface occurs. There is also preliminary evidence to suggest a role for PF08_0008 as an erythrocyte-binding molecule. The membrane anchorage of PF08_0008 has also been investigated, as it has sequence characteristics of both a type I membrane protein and a GPI-anchored protein. As part of this investigation, a line of parasites has been established in which endogenous PF08_0008 is expressed as a C-terminal fusion with a triple haemagglutinin epitope tag.

Index of Contents

Index of Figures

Abbreviations

A, C, G, T	adenine, cytosine, guanine, thymine
ASP	apical sushi protein
bp	base pairs
BSA	bovine serum albumin
cm^2	square centimetres
cm	Centimetres
CTD	C-terminal domain
DAPI	4',6-diamidino-2-phenylindole
DDT	dichloro-diphenyl-trichloroethane
DNA	deoxyribonucleic acid
dATP	deoxyadenosine triphosphate
dGTP	deoxyguanosine triphosphate
dH_2O	distilled H_2O
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
E. coli	*Escherichia coli*
EBA140/175/181	erythrocyte binding antigen 140 kDa/175 kDa/181 kDa
EDTA	ethylenediaminetetraacetic acid
EGTA	ethyleneglycoltetraacetic acid
g	grams
g	relative centrifugal force
GPI	glycophosphatidylinositol
HA	haemagglutinin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	horseradish peroxidase
IgG	immunoglobulin G
IPTG	isopropyl β-D-1-thiogalactopyranoside
J	joules
kb	kilobase pairs
kDa	kilodaltons
M	molar
mA	milliamperes
MES	2-(N-morpholino)ethanesulphonic acid
MBq	Megabecquerels

mg	milligrams
ml	millilitres
mM	millimolar
MOPS	3-(N-morpholino)propanesulphonic acid
MSP2/4/5/7/8/10	merozoite surface protein 2/4/5/7/8/10
ng	nanograms
Ni-NTA	nickel-nitrilotriacetic acid
nm	nanometres
NP40	nonidet P-40
NTD	N-terminal domain
OD	optical density
Py235	*Plasmodium yoelii* 235 kDa rhoptry protein
PBS	phosphate-buffered saline
pmol	picomoles
PTRAMP	*Plasmodium* thrombospondin-related apical merozoite protein
PvRBP1/2	*Plasmodium vivax* reticulocyte binding protein 1/2
RNA	ribonucleic acid
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SSC	sodium chloride-sodium citrate
STE	sodium chloride-tris-EDTA
TBE	tris-borate-EDTA
TEMED	tetramethylethylenediamine
U	units
µCi	microcuries
µg	micrograms
µl	microlitres
µm	micrometres
µM	micromolar
UV	Ultraviolet
V	volts
v/v	volume/volume
w/v	weight/volume

1. Introduction

1.1 Malaria: a global health problem

Malaria is a potentially fatal disease that still devastates poverty-stricken nations more than a century after the protozoan parasite *Plasmodium* was identified as its causative agent. *Plasmodium* is transmitted by a female mosquito of the genus *Anopheles* to a vertebrate host. Depending on the particular *Plasmodium* species, the vertebrate host can be a rodent, primate, bird or reptile.

The human malaria burden is not evenly distributed across the world - it falls primarily on the tropics, in particular South America, South and South-east Asia and much of Africa (Guerra et al., 2006). This is due to the incidence of *Plasmodium*-susceptible species of mosquito in these regions, which is largely attributable to climate.

Malaria was successfully eradicated from the northern parts of the United States and Southern Europe during the first half of the twentieth century, mainly through the destruction of mosquito breeding grounds, improved sanitation and some directed vector control (reviewed in Rieckmann, 2006). The use of the residual insecticide DDT eliminated malaria from the southern United States in the 1940s. Following the success of this program, in 1955 the World Health Organisation (WHO) commenced an ambitious malaria eradication program, aiming to eliminate the disease from many of the afflicted nations.

Amazingly, no attempts to eliminate malaria in sub-Saharan Africa were made, as it was considered an unfeasible proposition. Efforts in countries such as Sri Lanka and India were initially very successful, but these programs were unsustainable and ultimately failed. In 1972 the global eradication campaign was declared a failure by the WHO. This failure likely impacted the generosity of international donors and discouraged scientists. Indeed, of 1,223 new drugs developed between 1975 and 1996, only 3 were antimalarials (Trouiller & Olliaro, 1998), and so unsurprisingly there was little change in malaria mortality and morbidity during this time.

However, various social, political and environmental factors in combination with the emergence of insecticide-resistant mosquitoes and, most importantly, drug-resistant parasites, have exacerbated the already desperate situation, so that the malaria burden is now increasing (Greenwood & Mutabingwa, 2002).

An accurate estimation of mortality remains elusive, since many cases of malaria go undocumented. However, recent estimates suggest that just over 2.5 billion people live at risk of infection by *P. falciparum* (Guerra et al., 2006), the species of the parasite which is responsible for the vast majority of deaths. It is believed that there are in the region of 500 million clinical episodes of *falciparum* infection each year, with the vast majority of the estimated 1 million fatalities occurring in children under the age of five, in sub-Saharan Africa (Snow et al., 2005). The number of *P. vivax* infections, the most common *Plasmodium* infection found outside Africa, is thought to be in the region of 80 million per year (Mendis et al., 2001), with around 2.6 billion individuals currently living at risk (Guerra et al., 2006). Although *P. vivax* infections rarely cause death, they can result in a debilitating illness that can recur many years later.

In addition to the associated mortality, malaria also contributes to long-term health problems such as anaemia, premature births, developmental abnormalities, and deficiencies in the intellectual development of children. The direct healthcare costs of malaria morbidity, combined with its effect on productivity, means that there is also a massive economic burden on affected nations. Estimates suggest that the annual gross national product of malaria-endemic countries has risen 2% less than in malaria-free countries in the last 35 years (Sachs & Malaney, 2002), and that the disease costs Africa around $12 billion every year (WHO, 2003).

1.2 Life cycle of the parasite

Plasmodium belongs to the phylum Apicomplexa, which includes other parasites of human and veterinary importance such as *Toxoplasma*, *Theileria*, *Eimeria*, *Babesia*, and *Cryptosporidium*. Five species of *Plasmodium* are known to infect humans: *P. falciparum, P. vivax, P. ovale, P. malariae,* and *P. knowlesi.*

The life cycle of the parasite is complex and consists of both vertebrate and mosquito stages (figure 1.1). In the human host, infection begins with the

injection of infectious sporozoites into the body by a mosquito. Sporozoites are deposited in the dermis. Some will enter capillaries and travel to the liver, whereas others remain in the dermis or enter the lymphatic system (Amino et al., 2006). Sporozoites reaching the liver pass through the layer of sinusoidal cells lining the liver capillaries, before passing through a number of hepatocytes and ultimately coming to rest in a final liver parenchyma cell. Parasites then mature over a number of days in a membrane-bound compartment, differentiating into thousands of merozoite forms. It has recently been postulated that merozoite-filled vesicles or 'merosomes' bud off from the infected hepatocyte and are used to safely deliver the parasites to the liver sinusoid lumen, so that they can be released directly into the circulation, avoiding phagocytosis by resident Kupffer cells (Sturm et al., 2006). In *P. vivax* infections, liver stage parasites may remain dormant in the form of hypnozoites, and their activation can lead to a relapse years following the original infection.

Circulating merozoites released from the liver invade red blood cells and enter the asexual erythrocytic cycle. Parasites that have entered erythrocytes differentiate and multiply over a defined period of time, 48 hours in the case of *P. falciparum*, passing through various discrete developmental stages: ring stages, trophozoites and finally the schizont. Rupture of the schizont at the end of the 48 hour cycle results in the release of 16-32 merozoites into the circulation, which can then go on to invade a new erythrocyte, thus propagating the asexual blood stage of the parasite.

Some merozoites, upon invasion of a new red blood cell, will not replicate and will instead differentiate into male and female sexual forms in a process known as gametocytogenesis. Interestingly, it is thought that all individual merozoites from a single schizont are committed to asexual multiplication, female sexual development or male sexual development, before they undergo erythrocyte invasion (reviewed in (Talman et al., 2004)). Male microgametocytes and female macrogametocytes can be taken up into the midgut of the *Anopheline* vector if the mosquito takes a blood meal from an infected individual. Differentiation into microgametes and macrogametes then follows, with fertilisation of macrogamete by microgamete to form a zygote. A motile ookinete develops from the zygote, which actively invades the midgut epithelium and migrates to the extracellular space between the epithelium and the basal lamina of the gut. Here it comes to rest and develops into an oocyst. Sporozoites develop within the oocyst and are

released into the haemocoel, from which they can migrate to the salivary glands. Following invasion of the salivary gland epithelium, sporozoites are ready to infect another vertebrate.

Figure 1.1*: Life cycle of *Plasmodium falciparum

Figure from Bannister & Mitchell (2003).

(1) Infectious sporozoites are injected into the human host by a female *Anopheline* mosquito taking a blood meal. Sporozoites can enter a capillary and travel to the liver via the bloodstream, or they may remain in the dermis or enter the lymphatic system (Amino et al., 2006).

(2) Parasites reaching the liver first cross the layer of sinusoidal cells lining the liver capillaries before passing through a number of hepatocytes and ultimately coming to rest in a final liver parenchyma cell. Liver stage parasites differentiate and multiply over a period of around a week in a structure known as a hepatic schizont.

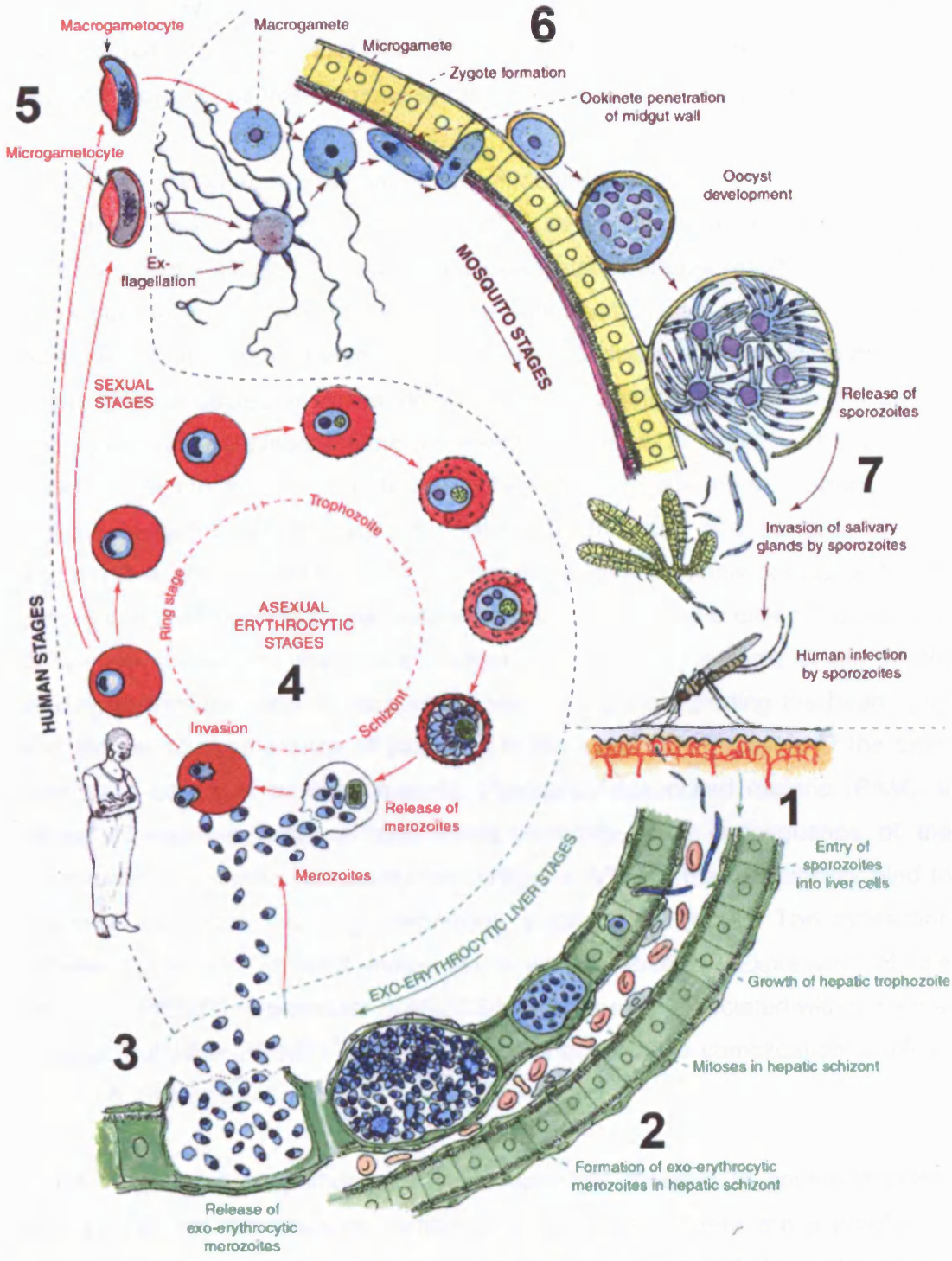
(3) Thousands of merozoites are delivered to the sinusoid lumen via merozoite-filled vesicles called 'merosomes' (Sturm et al., 2006) (not illustrated in this figure). This allows them to be released directly into the circulation, avoiding phagocytosis by resident Kupffer cells.

(4) Individual merozoites invade erythrocytes where they multiply asexually over a period of 48 hours, passing through discrete developmental stages such as ring stages, trophozoites and finally the schizont. Rupture of the schizont at the end of the 48 hour cycle results in the release of 16-32 merozoites, each of which can then invade an erythrocyte.

(5) Following erythrocyte invasion, some merozoites differentiate into microgametocytes or macrogametocytes, the male and female sexual forms of the parasite respectively.

(6) Gametocytes enter the midgut of the mosquito when a blood meal is taken from an infected individual. Differentiation into microgametes and macrogametes follows, with fertilisation of macrogamete by microgamete, to form a zygote. A motile ookinete develops from the zygote, which can migrate to the extracellular space between the mosquito midgut epithelium and the basal lamina of the gut. Here it comes to rest and develops into an oocyst. Sporozoites develop within the oocyst and are released into the haemocoele, from which they migrate to the salivary glands.

(7) After the invasion of the salivary gland epithelium, sporozoites are ready to infect another vertebrate as in (1).


Macrogametocyte
Macrogamete
Microgamete
6
Zygote formation
Ookinete penetration of midgut wall
5
Microgametocyte
Oocyst development
Ex-flagellation
MOSQUITO STAGES
SEXUAL STAGES
Release of sporozoites
7
Trophozoite
Invasion of salivary glands by sporozoites
Ring stage
ASEXUAL ERYTHROCYTIC STAGES
4
HUMAN STAGES
Human infection by sporozoites
Invasion
Schizont
Release of merozoites
Merozoites
1
Entry of sporozoites into liver cells
EXO-ERYTHROCYTIC LIVER STAGES
Growth of hepatic trophozoite
3
Mitoses in hepatic schizont
2
Formation of exo-erythrocytic merozoites in hepatic schizont
Release of exo-erythrocytic merozoites

1.3 Pathology of malaria

P. falciparum infections are responsible for the vast majority of malaria mortality, and it is the asexual blood stage of the parasite that is exclusively responsible for the clinical symptoms. Malaria can vary dramatically in its seriousness, from an asymptomatic infection, to a symptomatic febrile illness easily controlled by chemotherapy or the host immune system, to severe malaria and death.

Severe *P. falciparum* malaria is an ill-defined, extremely complex, multi-syndromic disorder. A common feature of severe malaria, and the principle determinant of survival, is metabolic acidosis which is often a result of a reduction in oxygen delivery to tissues. This oxygen deprivation can occur as a result of parasite-activated pro-inflammatory cytokine cascades, increased erythrocyte destruction, a decrease in the deformability of uninfected erythrocytes and a decrease in overall blood volume (reviewed in (Miller et al., 2002)). The ability of parasite-infected erythrocytes to cytoadhere to, and therefore obstruct, small blood vessels is also of great importance. Trophozoite- and schizont-infected erythrocytes are covered in knob-like protrusions, which enable contact with host cells (Luse & Miller, 1971), via a parasite-encoded molecule called *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) (Newbold et al., 1999), leading to sequestration of parasites in various organs including the heart, lung and kidney. Cytoadherence of parasites in the small blood vessels of the brain may be a cause of cerebral malaria. Pregnancy-associated malaria (PAM), a cause of massive maternal and infant morbidity, is a consequence of the expression of parasite variant surface antigens (VSAs), that specifically bind to placental receptors including chrondoitin sulphate A (CSA). The interaction between CSA and PfEMP1 molecules is well established. Expression of one variant of PfEMP1 in particular, VAR2CSA, is frequently associated with placental malaria, but other PfEMP1 variants also contribute to this complication (Duffy et al., 2006, Salanti et al., 2003).

A relatively small proportion of malaria episodes proceed to severe disease, around 1%, yet the absolute number is considerable. There are a number of parasite, host, geographic and social factors that can contribute to the clinical outcome of a malaria infection. For example, the drug resistance profile of the parasite, expression of parasite or host molecules important in cytoadherence, the presence of host immunity, and access to treatment.

1.4 History of malaria treatment

Quinine, an alkaloid extracted from the bark of the South American *Cinchona* tree, is the earliest known antimalarial to be used by Europeans and its use as a remedy is thought to date back to as early as the seventeenth century. Resochin, a quinine substitute, was developed in 1934 by a German scientist, Hans Andersag, and was at first disregarded as a candidate due to erroneous concerns over its toxicity. During Word War II US scientists obtained a supply of a German-developed compound called Sontochin, and following a few structural modifications chloroquine was born, although it was later discovered to be identical to the original Resochin.

Chloroquine was quickly identified as a cheap, safe and efficacious treatment for malaria, both radical cure and prophylaxis. However, in the late 1950s, within a decade of its introduction, chloroquine-resistant parasites were discovered in South America and South-east Asia and had reached Africa by the 1970s. The impact of chloroquine resistance on malaria mortality is astonishing, with a two-fold increase in the number of cases of severe malaria and death attributable to its emergence (Trape et al., 2002). Yet despite this, chloroquine still remains the first-line treatment in many countries - a fact that can be explained by the enormous disparity in the cost of chloroquine versus alternative antimalarials.

It is now generally accepted that combination therapy utilising drugs with different modes of action is the most suitable approach to malaria chemotherapy and prophylaxis. The plant *Artemisia annua* has been the source for the development of a new set of antimalarial compounds, and artemisinin-based combination therapies (ACT) are the regimen of choice of the WHO for the treatment of uncomplicated *P. falciparum* malaria.

1.5 Vaccine development

There are a variety of different types of malaria vaccine currently under development, targeting different aspects of the parasite biology (Matuschewski, 2006). Pre-erythrocytic vaccines aim to prevent hepatocyte invasion by sporozoites or to destroy infected cells, and erythrocytic stage vaccines are designed to eliminate, or limit the severity of, the blood-stage infection.

Transmission-blocking vaccines aim to prevent the spread of malaria, by preventing the infection of mosquitoes that feed on malarious individuals.

The development of an anti-cytoadhesion vaccine is theoretically possible, and is an attractive proposition. However, the cytoadherence-associated protein PfEMP1 is encoded by 60 members of the *var* multi-gene family. Any vaccine based on this molecule would therefore be very difficult to design, since *var* expression switching by the parasite allows evasion of the host immune system and a persistent infection can be maintained. Designing an anti-placental-adhesion vaccine on the other hand may be less problematic, due to the specific molecular nature of the interaction between parasite and host.

Antigens can be considered candidates for use in vaccines if they are essential for the propagation of the parasite life cycle, if they are associated with naturally-acquired immunity, or if they have been shown to be protective in animal models. Understandably, many of the parasite molecules thought to be involved in the invasion of host cells have been included in attempts to develop anti-malarial vaccines. These include thrombospondin-related anonymous protein (TRAP), which is known to be important in sporozoite invasion of hepatocytes, merozoite surface protein 1 (MSP1) and apical membrane antigen 1 (AMA1), which are both essential parasite molecules thought to be critical in merozoite invasion of erythrocytes.

It is likely that multi-component vaccines, which utilise more than one parasite antigen, will be important in overcoming the problems associated with antigen polymorphism. Since the vaccines that have entered clinical trials have so far failed to generate substantial, long-lived protection, it is obvious that considerable work will need to be done to reach the goal of an affordable, effective vaccine. In the meantime it will be useful to continue work to identify novel targets for further study.

1.6 An overview of the conserved molecular mechanism of *Plasmodium* host cell invasion

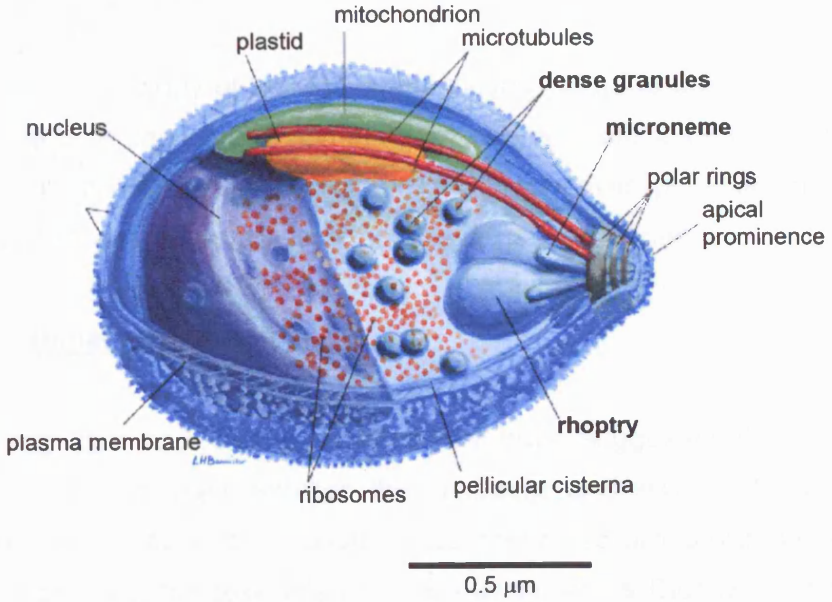
The *Plasmodium* developmental stages that are required to actively invade their host cell are the sporozoite, merozoite and ookinete. Although they are morphologically distinct, they possess certain shared characteristics. These

include a polarised morphology and a set of secretory organelles located at their apical end (figures 1.2a and 1.2b). The molecular mechanisms that drive motility and/or invasion of these zoites are also thought to be highly conserved (Baum et al., 2006).

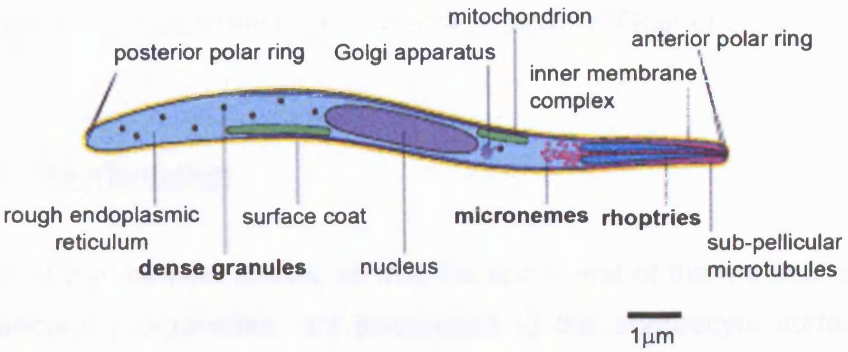
Zoites possess a membrane-anchored actin-myosin motor complex located between the parasite plasma membrane and the inner membrane complex (IMC), and homologues of the component molecules are found in all apicomplexan parasites. The motor consists of an unconventional class XIV myosin A (MyoA), which can interact with filamentous actin (F-actin) via its head domain. The actin-myosin motor is anchored to the outer membrane of the IMC by a complex consisting of MyoA tail-interacting protein (MTIP) (Bergman et al., 2003, Green et al., 2006b) and glideosome-associated proteins 45 and 50 (GAP45 and GAP50) (Baum et al., 2006).

Invasion is dependent on the regulated release of proteins from the apical organelles, including extracellular adhesin molecules that bind host cell receptors. The motor provides the driving force for movement, but in order for this action to be productive and forward-propelling, a link to the extracellular adhesins is necessary. In sporozoites an adhesin called thrombospondin-related anonymous protein (TRAP), a type I membrane protein, is critical. TRAP binds sulphated glucosaminoglycans on the liver cell surface, and importantly it is linked to the actin-myosin motor through its cytoplasmic tail, via an interaction with the F-actin-binding protein aldolase (Buscaglia et al., 2003). In ookinetes, a similar function is likely performed by the circumsporozoite/TRAP-related protein (CTRP) (Dessens et al., 1999). In merozoites, a third member of the TRAP family, MTRAP, is thought to function similarly (Baum et al., 2006) (see figure 1.2c). Although it has been shown to bind to aldolase *in vitro*, MTRAP has not been shown to have any erythrocyte binding capability. It is therefore possible that MTRAP binds to the erythrocyte surface indirectly via other parasite proteins, for example an erythrocyte binding antigen such as EBA175, and therefore acts as a link between extracellular adhesins and the motor. Force generated by the motor allows invasion to occur and the extracellular adhesins are simultaneously 'treadmilled' to the posterior of the parasite via their connection with the motor. The adhesins are ultimately cleaved by a protease, allowing the parasite to enter the host cell.

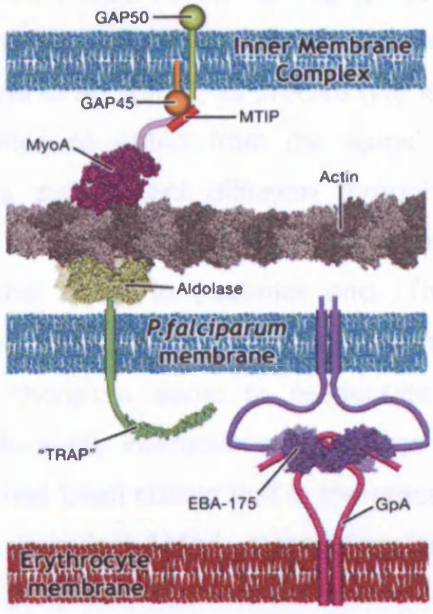
a)


mitochondrion
plastid
microtubules
dense granules
nucleus
microneme
polar rings
apical prominence
rhoptry
plasma membrane
ribosomes
pellicular cisterna
0.5 μm

b)


mitochondrion
posterior polar ring
Golgi apparatus
anterior polar ring
inner membrane complex
rough endoplasmic reticulum
surface coat
micronemes
rhoptries
sub-pellicular microtubules
dense granules
nucleus
1μm

c)


GAP50
Inner Membrane Complex
GAP45
MTIP
MyoA
Actin
Aldolase
P.falciparum membrane
"TRAP"
EBA-175
GpA
Erythrocyte membrane

1.7 Erythrocyte invasion

Since the human pathology of malaria is inextricably linked to the asexual blood stage of the parasite, there have been intense efforts to elucidate the molecular mechanisms of erythrocyte invasion in order to uncover potential drug and vaccine targets. However, the process remains largely undefined.

1.7.1 Initial attachment

Studies utilising light and electron microscopy have suggested that invasion begins with an initial contact between the merozoite and erythrocyte, which is thought to be a weak, reversible, low-affinity adherence occurring with a relatively large distance between the two cells (~12 nm) (Bannister & Dluzewski, 1990). It is thought that a large glycophosphatidylinositol (GPI)-anchored assembly of fragments of MSP1 and its associated proteins is responsible for the initial low-affinity interaction, although definitive evidence is lacking (Goel et al., 2003, Li et al., 2004).

1.7.2 Reorientation

Reorientation of the parasite occurs, so that the apical end of the merozoite, and hence the secretory organelles, are juxtaposed to the erythrocyte surface. A molecule that has been proposed to play a pivotal role in this stage is AMA1, a protein that is secreted onto the merozoite surface before invasion begins. Although it has been shown that AMA1 is highly conserved across the apicomplexan phylum, and is known to be an essential determinant of *P. falciparum* invasion (Triglia et al., 2000), its precise role is still unknown. It has been proposed that secretion of AMA1 from the apical organelles onto the merozoite surface, and its subsequent diffusion throughout the membrane, results in the creation of an AMA1 concentration gradient, with the amount present decreasing from the apical to posterior end (Thomas et al., 1990). Theoretically, any adhesive interactions occurring between AMA1 and the erythrocyte surface could therefore serve to reorientate the parasite to its required position, although such interactions have not been demonstrated conclusively. Additionally it has been shown that in the presence of a monoclonal antibody raised against *P. knowlesi* AMA1, merozoites fail to reorientate. The next step in the invasion process, which is the formation of a tight junction

between the two cells, also fails to occur in the presence of this antibody, strongly implicating AMA1 in either of those processes (Mitchell et al., 2004). In summary, the existing evidence suggests that AMA1 is likely to be a critical link between the initial weak interactions between the parasite and the erythrocyte, and the tighter, more specific interactions that follow.

1.7.3 Tight Junction Formation

Following reorientation, an irreversible tight junction forms between the apical prominence of the merozoite and the erythrocyte membrane, with a separation of around 4 nm (Bannister et al., 1986). This junction is composed of high-affinity interactions between parasite ligands released from the apical organelles, and erythrocyte receptors. Members of two *P. falciparum* protein families are likely to function as the parasite ligands at the tight junction. These are the Duffy-binding-like erythrocyte-binding protein (DBL-EBP), and *P. falciparum* reticulocyte-binding protein homologue (PfRH) families.

These proteins provide a variety of invasion pathways for the merozoite, as each protein interacts with a different erythrocyte receptor. Two members of the DBL-EBP family, EBA175 and EBA140, adhere to the sialic acid residues of glycophorin A (Sim et al., 1994) and glycophorin C (Maier et al., 2003) respectively, and a third member, EBA181 is believed to bind a different, but as yet undefined, trypsin-resistant erythrocyte receptor 'W' (Gilberger et al., 2003b). The N-terminal cysteine-rich domain of EBA175, region II, mediates binding to glycophorin A (Sim et al., 1994), and it is believed that similar domains in EBA140 and EBA181 perform similar functions.

P. vivax proteins PvRBP1 and PvRBP2 are specifically responsible for the reticulocyte-preference of this parasite (Galinski et al., 1992, Galinski et al., 2000), and homologues present in *P. yoelii*, the Py235 protein family, adhere to mouse erythrocytes, albeit through an undefined mechanism (Ogun & Holder, 1996). The PfRH family of proteins are homologues of the PvRBP and Py235 proteins and three of the six members are believed to play roles in invasion. PfRH1 binds to an undefined ligand 'Y' on erythrocytes (Rayner et al., 2001) and is essential for sialic acid-dependent invasion (Triglia et al., 2005). PfRH2b is essential for invasion via an unknown receptor 'Z' (Duraisingh et al., 2003b), although it has not been shown to have any binding activity. W2mef parasites are

able to switch from sialic acid-dependent to -independent invasion, by selection on neuraminidase-treated erythrocytes which lack sialoglycoproteins. This reversible switching has been shown to be correlated with, and dependent on, the expression of PfRH4 (Stubbs et al., 2005).

None of the genes encoding the DBL-EBP or PfRH proteins is essential for parasite growth, as they can all be successfully disrupted individually (Duraisingh et al., 2003a, Duraisingh et al., 2003b, Gilberger et al., 2003b, Reed et al., 2000, Stubbs et al., 2005, Triglia et al., 2005). This demonstrates the redundancy that exists within the process of invasion, and the ability of the parasite to invade via different pathways. This redundancy is highly beneficial to the parasite, since natural variation within the host means that parasites cannot rely on one particular invasion pathway. For example, an immune response may block a particular receptor-ligand interaction, and certain erythrocyte proteins and glycoconjugates are highly polymorphic within human populations, or may not always be present (Miller et al., 2002).

As mentioned previously, extracellular adhesins are required to connect with the actin-myosin motor for successful invasion to occur. The cytoplasmic tails of the DBL-EBP proteins do not seem to be directly linked to the actin-myosin motor (Gilberger et al., 2003a), and it is proposed that in merozoites a protein called MTRAP may form a link between the DBL-EBP and PfRH adhesins and the motor via aldolase, although this remains to be determined (Baum et al., 2006).

1.7.4 The moving junction

As invasion progresses, the ring-shaped tight junction established at the apex moves over the parasite towards the posterior end. The developing invagination of the host plasma membrane allows the parasite to enter its parasitophorous vacuole (PV). The moving junction itself is believed to be composed of a complex of parasite secretory organelle-derived molecules that function to exclude host cell membrane proteins from the nascent parasitophorous vacuole membrane (PVM) (Mordue et al., 1999). In *T. gondii*, the moving junction complex consists of TgAMA1 and at least one rhoptry neck protein, TgRON4 (Alexander et al., 2005). Similarly in *P. falciparum*, PfAMA1 co-immunoprecipitates with the *Plasmodium* homologue of TgRON4, suggesting a comparable complex exists in the malaria parasite (Alexander et al., 2006).

The removal or 'shedding' of the electron-dense coat from the surface of the merozoite is evident at the point of invasion (Aikawa et al., 1978, Bannister et al., 1975) and is likely to occur simultaneously with translocation of the moving junction over the surface of the parasite. Moreover, it has been shown that PfSUB2, a membrane-bound member of the subtilisin-like protease family, is responsible for the shedding of MSP1, the major merozoite surface coat constituent, as well as AMA1, which is also located on the surface of merozoites (Harris et al., 2005). Plasmodium thrombospondin-related apical merozoite protein (PTRAMP), another merozoite surface molecule, is a third substrate for PfSUB2, although the function of PTRAMP and the significance of its shedding is unknown (Green et al., 2006a).

PfSUB2 acts by cleavage at a juxtamembrane site within the ectodomain of its target molecule. It is proposed that PfSUB2 accumulates at the moving junction, shedding its target molecules as the parasite propels itself into the red blood cell, in a process described as 'shaving proteolysis'.

1.7.5 Disengagement of adhesive interactions

The current model for *Toxoplasma* tachyzoite invasion includes a shedding mechanism distinct from that of PfSUB2. Tachyzoite adhesive micronemal proteins (MICs) that bind to host cell ligands are translocated towards the rear of the parasite during invasion (Carruthers et al., 1999), via an interaction with the actin-myosin motor. The MICs come into contact with a recently characterised molecule called TgROM5, a rhomboid, at the posterior of the parasite during the final stages of invasion. Rhomboids are a class of protease that function by cleaving their substrate within the transmembrane domain and it has been shown that TgROM5 is responsible for the cleavage of MIC adhesins (Brossier et al., 2005).

Similarly, it has recently been shown that two *Plasmodium* rhomboid enzymes, PfROM1 and PfROM4, are responsible for the cleavage of a wide range of adhesins implicated in invasion, also within their transmembrane domains (Baker et al. 2006). For example, EBA175 is cleaved by PfROM4 in an event that is essential for parasite growth (O'Donnell et al., 2006).

In summary, it appears that during invasion, *Plasmodium* proteases act through two distinct mechanisms, but in a coordinated fashion, in order to allow the progression of the parasite into its vacuole within the host cell:

- PfSUB2 removes the merozoite surface coat at the moving junction,
- PfROM1 and PfROM4 disengage adhesive interactions between erythrocyte ligands and parasite receptors.

1.8 Apical organelles and their complement

The apical organelles of the invasive sporozoite and merozoite consist of rhoptries, micronemes, and dense granules. Rhoptries are paired club-shaped vesicles, whereas micronemes are more numerous with up to 40 per cell. Micronemes are smaller in size than rhoptries and are clustered around the rhoptry ducts. Importantly, these vesicles secrete their contents from the zoite apex at the point of host cell invasion. Dense granules (DG) are microspheres with a diameter of approximately 200 nm and secretion of DG-resident proteins is thought to occur at a sub-apical location, following the fusion of DG and parasite plasma membranes.

It has been shown in *Toxoplasma* that there is a sequential secretion of proteins from the different sets of organelles: microneme contents are secreted initially, followed by the rhoptry- and ultimately dense granule-resident proteins (Carruthers & Sibley, 1997). This led to the suggestion that proteins are functionally segregated: micronemal proteins are either adhesins or molecules that link the parasite surface to the actin-myosin motor, rhoptry proteins are responsible for the formation of the PVM, and dense granule proteins are involved in host cell modification.

Although this suggestion may be broadly accurate, and is supported by evidence such as the absence of rhoptries in ookinetes, which do not utilise a PV, it is likely that there is significant overlap in the roles played by proteins from these organelles (Kats et al., 2006). For example, although the DBL-EBP family of adhesive proteins are microneme-resident, the rhoptry neck protein PfRH1 is also thought to act as an adhesin through binding to an undefined glycosylated erythrocyte receptor (Rayner et al., 2001).

1.8.1 Microneme proteins of *P. falciparum*

The best characterised of the microneme proteins are the DBL-EBP family of proteins and AMA1, whose roles have already been discussed. MTRAP, which is potentially the link between the merozoite motor complex and extracellular adhesins, is also micronemal, along with the merozoite 'sheddase' PfSUB2. Apical sushi protein (ASP) is a recently described GPI-anchored protein, which has also been localised to the micronemes. ASP has a sushi-domain, which is known to be involved in protein-protein or protein-ligand interactions in other species (O'Keeffe et al., 2005). However, a specific role for this protein has yet to be defined.

Micronemal proteins are presumed to be trafficked initially via a classical N-terminal signal sequence that allows them to cross the endoplasmic reticulum and enter into the secretory pathway. However, the details of specific trafficking to the micronemes are yet to be fully elucidated. The DBL-EBP proteins are single span transmembrane proteins with a cytoplasmic domain and ectodomain. It has recently been shown that a C-terminal cysteine-rich region of the ectodomain of these proteins is responsible for targeting to the micronemes (Treeck et al., 2006). It seems likely that other micronemal proteins that do not possess this domain are trafficked by alternative mechanisms.

1.8.2 Processing and shedding of microneme proteins

The proteolytic processing of apicomplexan microneme proteins is widespread, and is often essential for their correct functioning. In *Toxoplasma gondii*, microneme proteins exist as complexes, which are ultimately released onto the parasite surface following a rise in intracellular calcium levels (Carruthers & Sibley, 1999). Most of these proteins are subject to proteolytic modifications following their entry into the secretory pathway (Dowse & Soldati, 2004). For example, the TgMIC2/TgM2AP complex is proteolytically processed and ultimately shed in a process involving three distinct microneme proteases that is essential for successful invasion (Brossier et al., 2003, Zhou et al., 2004).

This phenomenon also occurs in *Plasmodium*, with AMA1 processing widely studied and shown to be a multi-step process essential for successful invasion. The initial processing of PfAMA1 occurs shortly after its trafficking to the

micronemes, by way of an N-terminal cleavage event (Howell et al., 2001, Narum & Thomas, 1994). Other micronemal proteins that appear to be processed in a similar manner include MTRAP (Baum et al., 2006) and ASP (O'Keeffe et al., 2005). Some micronemal proteins are subject to further proteolysis, which results in them being shed into culture supernatant. For example, the N-terminally truncated form of AMA1, which is transferred to the parasite surface, is ultimately shed in two alternative forms, following further proteolysis by PfSUB2 (Harris et al., 2005, Howell et al., 2003). Interestingly, PfSUB2 is also responsible for the shedding of MSP1 (Harris et al., 2005) and PTRAMP into culture supernatant (Green et al., 2006a).

It has been shown that two of these shedding events are critical for invasion. Anti-AMA1 antibodies that inhibit shedding can also inhibit invasion (Dutta et al., 2005, Dutta et al., 2003). It has also been shown that the shedding of EBA175 is essential for propagation of the asexual blood stage (O'Donnell et al., 2006).

Since shedding of AMA1 occurs predominantly via a juxtamembrane cleavage just 29 residues away from the predicted transmembrane region (Howell et al., 2005, Howell et al., 2003), a short 'stub' of AMA1 ectodomain is generated, which remains anchored to the parasite surface via its intact transmembrane domain. It has been shown that this stub is carried into newly invaded ring stage parasites, where it remains associated with the parasite plasma membrane (Howell et al., 2005). This is reminiscent of the fate of $MSP1_{19}$, the residual stub of MSP1, which has also been identified in ring stage parasites (Blackman et al., 1990) and localises to the food vacuole (Ling et al, manuscript in preparation). Whether these residual stubs serve any function in the ring stage parasite remains to be determined.

1.9 GPI-anchored proteins of *P. falciparum*

1.9.1 The GPI-anchor and its addition

The addition of a GPI-anchor to proteins is a widespread phenomenon in both lower and higher eukaryotes (Eisenhaber et al., 2001), including *Trypanosoma* and *Plasmodium* parasites, and serves to anchor proteins in the plasma membrane of a cell. The GPI moiety itself is a complex glycolipid and its structure is shown in figure 1.3a.

Addition of a GPI-anchor requires the presence of an N-terminal signal sequence, which allows co-translational insertion of the polypeptide into the endoplasmic reticulum (ER). Before anchorage can occur, the entire polypeptide chain is translocated across the ER membrane and into the lumen (Dalley & Bulleid, 2003a). Once in the lumen, a C-terminal GPI-anchor signal sequence is recognised by a transamidase complex.

The GPI-anchor signal sequence consists of four components, detailed in figure 1.3b (Eisenhaber et al., 2001): briefly, a C-terminal hydrophobic region is preceded by a short hydrophilic spacer that is linked to the anchor addition site (residue ω). The transamidase complex catalyses the concomitant removal of the GPI-anchor signal sequence, by way of a proteolytic cleavage between residues ω and ω+1, and its replacement with a pre-synthesised, membrane-bound GPI-anchor (figure 1.3c). The protein is thus anchored, and is exposed initially to the luminal side of the endoplasmic reticulum, and ultimately the cell exterior.

Single-pass transmembrane proteins are thought to be anchored in a similar but distinct way. The transmembrane domain itself consists of a hydrophobic region followed by a 'stop transfer' sequence that consists of charged residues. The protein is co-translationally passed through the translocon pore of the ER membrane and into the lumen, until the ribosome encounters and recognises the 'stop transfer' sequence (High & Dobberstein, 1992). This triggers the blockage of the translocon pore on the luminal side, and lateral translocation of the polypeptide within the ER membrane occurs, resulting in the integration of the transmembrane domain into the lipid bilayer (Martoglio et al., 1995).

The transmembrane and GPI-anchor signal sequences both contain hydrophobic regions, and there are characteristics shared between the two. It is likely that subtle differences in the hydrophobic region of a GPI-addition sequence prevent its recognition and lateral translocation within the ER membrane, as if it were a transmembrane domain (Dalley & Bulleid, 2003b). Nonetheless, one might expect GPI-addition sequences to be wrongly assigned as transmembrane domains by algorithms. For example, bioinformatic analysis of the ASP amino acid sequence predicted a C-terminal transmembrane anchor (O'Keeffe et al., 2005), although it was later experimentally shown to be GPI-anchored (Gilson et al., 2006).

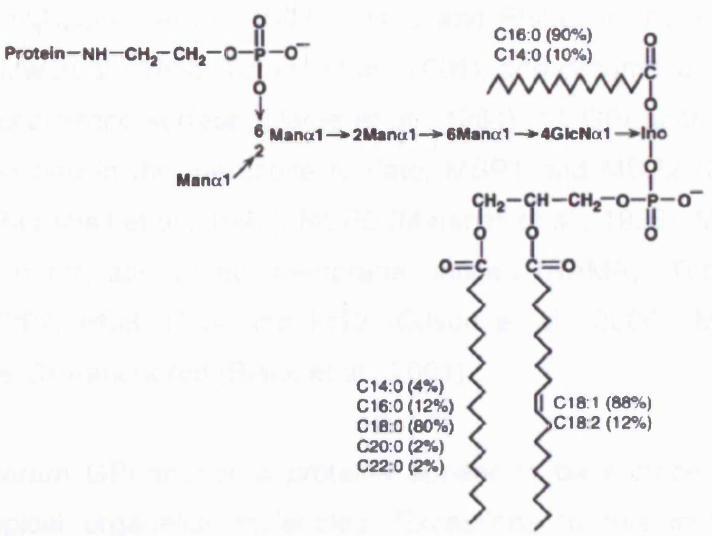
Figure 1.3*: GPI-anchor structure and addition in *P. falciparum

a) The presumed structure of the *P. falciparum* GPI-anchor. Diagram from Naik *et al.* (2000). Molar proportions of the three fatty acids are indicated. Two fatty acids within a hydrophobic phosphatidylinositol group act as the membrane anchor. A carbohydrate linker consisting of mannose and glucosamine connects the phosphatidylinositol group to the C-terminal residue of the anchored protein via a phosphoethanolamine residue.

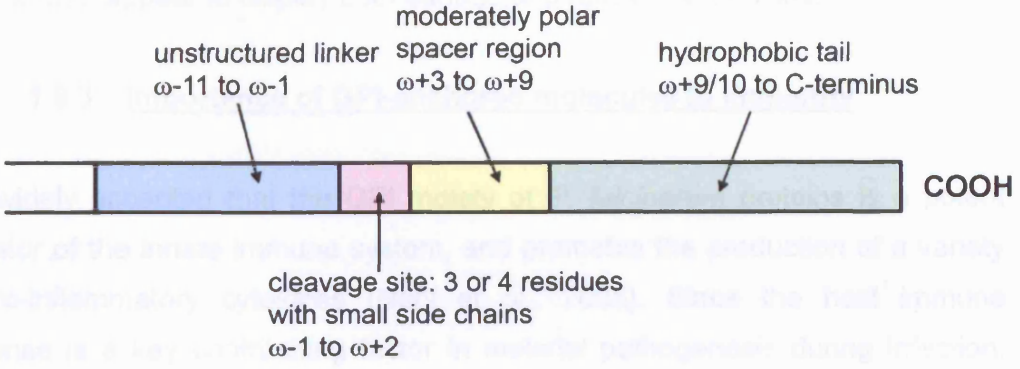
b) The sequence requirements for GPI-anchor addition. Cleavage of the signal sequence occurs between residues ω and ω+1. The cleavage site itself (coloured pink, residues ω-1 to ω+2) consists of amino acid residues with small side chains, and is followed by a moderately polar spacer region (yellow, residues ω+3 to ω+9). A hydrophobic tail is found approximately 9 or 10 residues C-terminal of the ω site (green), and extends to the C-terminus of the protein. An unstructured linker is found from residues ω-11 to ω-1 (blue).

c) Diagram from Dalley & Bulleid (2003a). The entire polypeptide chain of a GPI-anchored protein is fully translocated into the lumen of the endoplasmic reticulum, prior to anchor addition. The GPI signal sequence is recognised by a membrane-bound transamidase complex that catalyses the removal of the GPI signal sequence and its replacement with a pre-synthesised GPI-anchor.

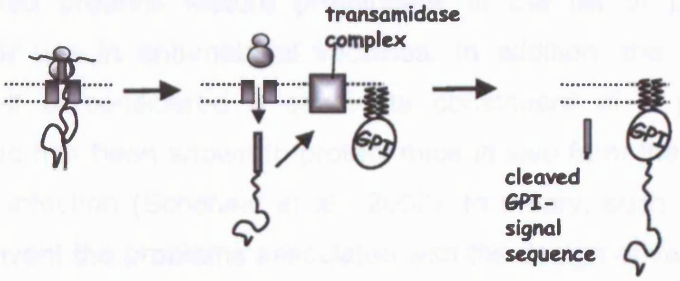
a)


Protein—NH—CH2—CH2—O—P—O⁻
C16:0 (90%)
C14:0 (10%)
6
Manα1→2Manα1→6Manα1→4GlcNα1→Ino
2
Manα1
CH2—CH—CH2—O—P—O⁻
O=C
C=O
C14:0 (4%)
C16:0 (12%)
C18:0 (80%)
C20:0 (2%)
C22:0 (2%)
C18:1 (88%)
C18:2 (12%)

b)


unstructured linker
ω-11 to ω-1
moderately polar
spacer region
ω+3 to ω+9
hydrophobic tail
ω+9/10 to C-terminus
COOH
cleavage site: 3 or 4 residues
with small side chains
ω-1 to ω+2

c)


transamidase
complex
GPI
GPI
cleaved
GPI-
signal
sequence

1.9.2 GPI-anchored proteins in *Plasmodium*

All of the extracellular forms of *P. falciparum* are covered in proteins that are either known or presumed to be GPI-anchored. For example Pfs48/45 on the gamete surface (Kocken et al., 1993), Pfs25 and Pfs28 on the surface of the ookinete (Kaslow et al., 1988, Tomas et al., 2001), and circumsporozoite protein (CSP) on the sporozoite surface (Dame et al., 1984). 11 GPI-anchored proteins have been identified in the merozoite to date, MSP1 and MSP2 (Gerold et al., 1996), MSP4 (Marshall et al., 1997), MSP5 (Marshall et al., 1998), MSP10 (Black et al., 2003), rhoptry-associated membrane antigen (RAMA) (Topolska et al., 2004), ASP, Pf92, Pf38, Pf34 and Pf12 (Gilson et al., 2006). MSP8 is also presumed to be GPI-anchored (Black et al., 2001).

Most *P. falciparum* GPI-anchored proteins appear to be surface antigens, as opposed to apical organellar molecules. Exceptions to this include rhoptry-associated membrane antigen (RAMA) (Topolska et al., 2004), and ASP, a microneme protein. Interestingly, Pf38 (Sanders et al., 2005) and MSP10 (Black et al., 2003), appear to display both surface and apical distributions.

1.9.3 Importance of GPI-anchored molecules in immunity

It is widely accepted that the GPI moiety of *P. falciparum* proteins is a potent activator of the innate immune system, and promotes the production of a variety of pro-inflammatory cytokines (Nebl et al., 2005). Since the host immune response is a key contributing factor in malarial pathogenesis during infection, GPI-anchored proteins are regarded as critical toxins.

GPI-anchored proteins feature prominently in the list of potential candidate antigens for use in anti-malarial vaccines. In addition, the *P. falciparum* GPI moiety itself is considered a candidate constituent of a potential anti-toxin vaccine, and has been shown to protect mice *in vivo* from the pathogenesis of a *P. berghei* infection (Schofield et al., 2002). In theory, such anti-toxin vaccines may circumvent the problems associated with the design of vaccines designed to neutralise parasite proteins. These problems include antigenic variation, functional redundancy, and the time taken by the host to acquire immunity.

1.10 The post-genomic era of research

The fruitless efforts to develop a licensed malaria vaccine, along with the emergence of drug-resistant parasites and insecticide-resistant mosquitoes highlight the urgency with which new points of attack to combat malaria need to be identified. The arrival of the *P. falciparum* genome sequence (Gardner et al., 2002), along with its transcription (Bozdech et al., 2003, Le Roch et al., 2003) and proteomic (Florens et al., 2002, Lasonder et al., 2002) profiles, has provided scientists with a great opportunity to identify novel drug and vaccine candidates.

1.10.1 Transcriptome data

Two independent DNA microarray studies were published in 2003 and have provided invaluable information with regards to the expression levels of genes throughout the intra-erythrocytic blood cycle and other developmental stages. One of the main findings of these two studies is that groups of genes with similar levels and temporal patterns of expression can be assembled into 'clusters', with each member of a particular cluster proposed to be functionally related.

The Bozdech *et al.* study (2003), which utilised asexual stages of *P. falciparum* clone HB3, revealed that the majority of intra-erythrocytic expression profiles show some degree of periodicity, with a maximum and minimum level of transcript abundance at defined points within the 48 hour cycle. Few genes are transcribed at a steady level throughout. Phaseograms show a cascade of continuous transcription throughout the cycle, with no clearly defined boundaries, and in general, timing of gene expression is appropriate for the function of the resultant protein. For example, genes involved in DNA replication and synthesis are maximally expressed at the early-schizont stage, when they would be required by the parasite. Additionally, the expression of all genes presumed to be involved in invasion, for example, AMA1, MSP1 and EBA175, is sharply induced to a large amplitude in mid- to late-schizont stages. This cascade of expression profiles is likely to correspond to a programmed cascade of cellular processes and allow the necessary molecules to be present at the appropriate time to fulfil their role in the parasite.

The Le Roch *et al.* study (2003) utilised RNA from sporozoite, merozoite and gametocyte stages, in addition to six asexual blood stages, from *P. falciparum*

3D7 parasites. By grouping genes together on the basis of their time of expression in the life cycle using a *k*-means clustering algorithm, the authors were able to demonstrate that membership in a particular cluster was highly predictive of function, based on the distribution of proteins with known functions. For instance, genes assigned to cluster 15 were highly expressed in the schizont stage and representative members included *AMA1*, *MSP1*, *EBA140* and *PfSUB2*, suggesting that members of this cluster may function in invasion.

A further transcriptome study covering the *P. falciparum* sporozoite, asexual intra-erythrocyctic and gametocyte stages was published in 2005 and used an algorithm called ontology-based pattern identification (OPI) (Young *et al.*, 2005). This algorithm utilises gene annotations as well as temporal expression patterns to generate clusters of genes, with a view to enabling the functions of hypothetical genes to be predicted.

1.10.2 Proteome data

Several high-throughput mass spectrometry-based studies of the proteome of *P. falciparum* have been published (Florens et al., 2002, Lasonder et al., 2002), and importantly the results appear to correlate with the transcriptome data described above.

The Florens *et al.* (2002) study utilised multi-dimensional protein identification technology (MudPIT), a technique that combines high-resolution liquid chromatography and tandem mass spectrometry, and identified around 2400 parasite proteins, nearly half of all predicted gene products, from sporozoites, merozoites, trophozoites and gametocytes. The Lasonder *et al.* (2002) mass spectrometric proteome study was able to confidently identify 1289 parasite proteins from asexual blood stages, gametocytes and gametes.

However, there are some discrepancies between transcriptome and proteome data. For example, >50% of proteins appeared stage specific in the Florens *et al.* (2002) study, whereas <30% of transcripts in the Le Roch *et al.* (2003) study were described as being unique to one stage. Also, there are significant temporal discrepancies between mRNA and protein abundance in *P. falciparum* (Le Roch et al., 2004). However these two inconsistencies can be explained by the differing sensitivities of the transcriptome and proteome methods, and the possibility of

post-transcriptional control respectively. There is no doubt that these data still remain invaluable resources in their own right. The existence of such data sets also makes it relatively easy to identify proteins found in particular stages of interest, for example those only expressed in sporozoites and merozoites.

1.10.3 The 'GPI proteome'

A couple of recent studies have been important in the attempt to generate a GPI-anchored proteome of *P. falciparum*. Proteome analysis of purified detergent-resistant membrane (DRM) domains from *P. falciparum* schizonts, which are highly enriched in GPI-anchored proteins, enabled the identification of novel merozoite surface-resident, putative GPI-anchored proteins such as Pf38 and Pf12 (Sanders et al., 2005). A different approach which employed radiolabelled GPI precursors was also able to identify novel GPI-anchored proteins, including the non-surface-resident proteins RAMA and ASP (Gilson et al., 2006).

It is highly likely that there are additional, unidentified GPI-anchored proteins present in schizonts and merozoites. The authors of the Gilson *et al.* (2006) study estimated that MSP1 and MSP2 constitute around two-thirds of the merozoite surface-resident GPI-anchored proteins, and therefore identifying proteins present at a much lower copy number is a task which may prove difficult.

1.11 Aims of this project

The present study aims to identify and characterise a novel *P. falciparum* protein that is involved in the invasion of erythrocytes.

Functional genomics has effectively revolutionised the manner in which potentially important genes, in particular, those implicated in host cell invasion, are identified and studied. Comparison of the *P. falciparum* nuclear genome with other genomes revealed that around 60% of the ~5300 predicted genes encode proteins with no assigned function (Gardner et al., 2002). Consequently, there is a formidable number of novel potential drug and vaccine targets to contend with. It is therefore necessary to make efficient use of the wealth of data available and refine the list of potential novel targets by considering some important points.

For instance, it is worth thinking about the common features shared by the molecules that are already known or are highly likely to be involved in the invasion process. These include:

- The presence of a signal peptide, indicating transport through the secretory pathway, possibly to apical organelles or the parasite surface
- The presence of a single transmembrane domain, as exemplified by AMA1, and the PfRH and DBL-EBP protein families
- A late blood stage expression profile, suggestive of a role in erythrocyte invasion

Sporozoites and merozoites are both actively invasive, and although the cells they invade may differ, the basic machinery of invasion is conserved between them (Baum et al., 2006). Therefore it should be considered that genes expressed in both stages could be critical in host cell invasion. Indeed, it has been shown that AMA1, which is expressed and critical in the asexual blood stage, is also expressed in sporozoites and has a critical role in the invasion of the hepatocyte (Silvie et al., 2004).

The main aims of this project are:

- To utilise published proteome, genome and transcriptome data to identify by bioinformatic means, a gene of interest for study as detailed below:
    1. Published proteomic data (Florens et al., 2002) will be used initially to identify uncharacterised *P. falciparum* proteins expressed only in merozoite and sporozoite fractions that possess a predicted single transmembrane domain and signal peptide.
    2. The amino acid sequences of candidate proteins will be analysed to identify any conserved domains or features of interest, as well as orthologues that may be present in other *Plasmodium* or apicomplexan species.
    3. Transcriptome data for the genes encoding the identified candidate proteins will be retrieved and analysed to determine if any genes display a late blood stage expression profile.
    4. A decision on the choice of gene for study will be taken based on the outcome of performing steps 1-3.
- To characterise the chosen gene and its corresponding protein with respect to:
    - location within the parasite
    - proteolytic processing, if any
    - potential role in invasion, if any.

2. Materials and Methods

2.1 General Materials

2.1.1. Chemicals

All chemicals and ingredients used throughout the project were obtained from Sigma unless stated otherwise.

2.1.2 Solutions

Commonly used solutions included:

PBS:	137 mM NaCl, 3 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4 in dH_2O
1 x TBE:	90 mM Tris-Borate, 2 mM EDTA in dH_2O, *p*H 8.0
Luria-Bertani broth (LB):	1% w/v bacto-tryptone, 0.5% w/v bacto-yeast extract, 170 mM NaCl in dH_2O, autoclaved
LB-Amp:	As LB, with 100 μg ml^{-1} ampicillin
LB-Kan:	As LB, with 25 μg ml^{-1} kanamycin
LB-Agar:	As LB with 1.5% w/v agar
LB-Amp/Agar	LB-Agar prepared as above, autoclaved and cooled before the addition of 100 μg ml^{-1} ampicillin
LB Kan/Agar	LB-Agar prepared as above, autoclaved and cooled before the addition of 25 μg ml^{-1} kanamycin

2.1.3 Centrifuges

E. coli cultures were spun down in disposable polypropylene centrifuge tubes (Corning) in a Beckman J-6B refrigerated centrifuge with rotor JS5.2. Bacterial protein extraction centrifugation steps were carried out in a Beckman J2-21 refrigerated centrifuge with a JA-20 rotor and reusable polypropylene Oak Ridge centrifuge tubes (Nalgene).

For phenol/chloroform extraction of DNA (cf. 2.5.1), precipitated DNA was pelleted in Corex tubes by centrifugation in a Beckman J2-21 refrigerated centrifuge with a JA-20 rotor.

For centrifugation of small volumes (≤2 ml), a Heraeus Biofuge Pico microcentrifuge was used, with rotor #3325B, unless refrigeration was required, in which case a Heraeus Biofuge Fresco microcentrifuge was used with rotor #3328.

For the buffer exchange and/or concentration of solutions >2 ml, a Heraeus PrimoR refrigerated centrifuge was used with rotor #7591 and 20 ml Vivaspin centrifugal concentrators with an appropriate molecular weight cut-off (Sartorius). For buffer exchange and/or concentration of small volumes (≤2 ml), 500 µl Vivaspin centrifugal concentrators with an appropriate molecular weight cut-off were used with either one of the microcentrifuges detailed above.

For the extraction of membrane-associated proteins, centrifugation steps were carried out in a Beckman TL-100 ultracentrifuge, with a TLA-100.3 rotor, using 1.5 ml polyallomer microfuge tubes (Beckman).

When handling parasites, a Sigma 4-10 centrifuge was used with rotor #11140. For smaller volumes a Heraues Biofuge Pico microcentrifuge was used, with rotor #3325B.

2.2 Bioinformatics

2.2.1 Identification of a candidate gene for study and initial bioinformatics

Data from the Florens *et al.* (2002) proteomic study, which identified 2415 proteins from four developmental stages of *P. falciparum*, were used in order to select an uncharacterised *P. falciparum* gene for study. A set of criteria was designed in order to identify proteins from this dataset that may play a role in host cell invasion:

- Proteins must be predicted to contain a signal peptide
- Proteins must be predicted to have a single transmembrane domain

- Proteins must have been detected in merozoite and sporozoites fractions only.

The dataset was sorted based on the criteria above, and this resulted in the identification of *PF08_0008* as a suitable gene for study. Information on *PF08_0008* was gathered from the *Plasmodium* genome resource PlasmoDB (http://plasmodb.org), as well as a number of published transcriptome studies (Bozdech et al., 2003, Le Roch et al., 2003, Llinas et al., 2006).

The biological sequence alignment editor BioEdit (Tom Hall, Ibis Therapeutics) was used to create a hydrophobicity plot for PF08_0008, using the Kyte & Doolittle (1982) method, with a window size of 13.

A number of algorithms were used to predict any transmembrane anchorage of the chosen protein: TMHMM2 (http://www.cbs.dtu.dk/services/, (Krogh et al., 2001)), TOPPRED2 (http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html, (Claros & von Heijne, 1994)) and TMPRED (http://ch.embnet.org, (Hofmann & Stoffel, 1993)).

Two web-based GPI prediction servers were used to determine the likelihood of PF08_0008 being a GPI-anchored protein: the Big-π GPI prediction server (http://mendel.imp.ac.at/gpi/gpi_server.html, (Eisenhaber et al., 1999)) and DGPI (http://129.194.185.165/dgpi/DGPI_demo_en.html, (Kronegg & Buloz, 1999)). The PSIPred tool (bioinf.cs.ucl.ac.uk/psipred, (McGuffin et al., 2000)) was used to predict secondary structure in the PF08_0008 C-terminus.

The PlasmoAP tool (http://plasmoDB.org, (Foth et al., 2003)) was used in order to predict the likelihood that PF08_0008 is targeted to the parasite apicoplast. The SignalP server (http://www.cbs.dtu.dk/services/SignalP, (Nielsen et al., 1997)) was used to predict the presence of a signal peptide at the N-terminus of PF08_0008.

2.2.2 Database searching and sequence alignment

Searches for similar protein sequences were carried out using the basic local alignment search tool (BLAST) (Altschul et al., 1990), available at the National Centre for BioInformatics (http://ncbi.nlm.nih.gov/BLAST/), with default search

parameters. Sequences of relevant hits from *Plasmodium* species were downloaded from PlasmoDB and were aligned using the BioEdit sequence alignment software. Multiple alignments were carried out using ClustalW (Thompson et al., 1994).

The Simple Modular Architecture Research Tool (SMART) (http://smart.embl-heidelberg.de/, (Schultz et al., 1998)) and ProSite (http://expasy.ch/prosite (Hulo et al., 2006)) searchable databases were used to identify conserved protein domains in PF08_0008.

2.3 *Plasmodium falciparum* culture

All solutions were pre-warmed to 37°C before use. The magnet-purification of schizonts, purification of merozoites, preparation of culture supernatant and metabolic labelling of parasites were carried out by Muni Grainger (NIMR, Division of Parasitology).

2.3.1 Parasites and media

RPMI w/o:	RPMI 1640 enriched with 27 mM $NaHCO_3$, 25 mM HEPES, 25 µg ml^{-1} gentamycin, 50 µg ml^{-1} hypoxanthine and 2 mg ml^{-1} glucose (Gibco)
RPMI w/Albumax:	As above supplemented with 5 mg ml^{-1} AlbuMAX (Gibco)

Human erythrocytes (typically O+) were obtained from the National Blood Transfusion Service and were stored at 4°C, for no longer than 1 month. Cells were washed three times in RPMI w/o before use, with centrifugation at 2700rpm/1483*g* for 10 minutes.

2.3.2 Continuous culture

Phosphate buffer: 700 mM Na_2HPO_4, 515 mM KH_2PO_4, *p*H 7.2

P. falciparum asexual stages were cultured in human erythrocytes and RPMI w/Albumax. Typically a 175 cm^2 flask was used, with a culture volume not exceeding 100 ml. Parasitemia was maintained at or below 15% and the haematocrit used was typically 1-2%. Cultures were gassed with a mixture of 7%

CO_2, 5% O_2, and 88% N_2 and incubated at 37°C. Culture medium was replaced typically every 2 days. Cultures were harvested by centrifugation at 1600rpm/521*g* for 8 minutes, the old medium aspirated off and fresh medium added.

Cultures were analysed by staining thin blood smears. Smears were briefly fixed with methanol. Stain was made up by diluting Giemsa's staining solution (BDH) 1 in 5 in phosphate buffer, then passing through a 0.2 µm filter. Smears were stained for 5 minutes then rinsed with water, dried and viewed by light microscopy.

2.3.3 Synchronisation of parasites and purification of schizonts

Isotonic Percoll:	1 volume 10 x PBS, 9 volumes Percoll (GE Healthcare)
70% isotonic Percoll:	7 volumes isotonic Percoll, 3 volumes RPMI w/o. 0.2 µm filtered and stored at 4°C.
Sorbitol:	5% w/v sorbitol in PBS, 0.2 µm filtered and stored at 4°C.

A crude synchronisation was achieved by centrifuging the culture at 1600rpm/521*g* for 8 minutes, then resuspending the pellet in 5 volumes of sorbitol. Following a 10 minute incubation at room temperature, cells were centrifuged at 1300rpm/344*g* for 3 minutes. Sorbitol was aspirated and cells were resuspended in fresh RPMI w/o and centrifuged for a further 3 minutes at 1300rpm/344*g*. Medium was aspirated and cells returned to culture.

Cultures were tightly synchronised using a two-stage approach. A late blood-stage culture was centrifuged at 1600rpm/521*g* for 8 minutes, the pellet was resuspended in 5 volumes of RPMI w/o and layered on top of 10 ml 70% isotonic Percoll. Following centrifugation at 2200rpm/986*g* for 11 minutes, the layer of schizonts was removed and resuspended in RPMI w/o. Schizonts were centrifuged for a further 5 minutes at 1800rpm/660*g*. Medium was aspirated off and schizonts were returned to culture for a period of 1-4 hours. Remaining schizonts were lysed with sorbitol treatment.

A MACS type D depletion column was used in combination with a SuperMacs II magnetic separator (Miltenyi Biotec) to purify schizonts from uninfected erythrocytes. The column was washed with ethanol, followed by dH_2O, and was

equilibrated with RPMI w/Albumax. Infected erythrocytes from late blood-stage *P. falciparum* cultures were harvested by centrifugation and re-suspended in an equal volume of RPMI w/Albumax. The cell suspension was then passed through the MACS column under gravity flow. The column was washed with four column volumes of RPMI w/Albumax until no erythrocytes could be seen in the flow-through. The column was removed from the magnetic separator and washed with 50 ml RPMI w/Albumax in order to allow collection of schizonts. If required for protein analysis, schizont material was harvested by centrifugation, the medium aspirated off and the cellular material stored at -70°C until used.

Parasites to be used in invasion inhibition assays were synchronised using a two-stage magnet purification approach. Late blood-stage parasites were purified from culture using the MACS depletion column as described above, and returned to culture for 2 hours. Remaining schizonts were then removed from the culture using the MACS depletion column, and the flow-through containing newly-invaded ring stages and uninfected erythrocytes was retained and returned to culture.

2.3.4 Metabolic labelling of parasites

Tightly synchronised late blood-stage schizonts were magnet purified as in 2.3.3, then resuspended to a 2-3% haematocrit in RPMI w/Albumax, free of methionine and cysteine, except for a supplement of 100 µCi ml^{-1} [^{35}S]methionine and [^{35}S]cysteine (Pro-Mix, GE Healthcare). Parasites were cultured in the presence of the radiolabel for 2 hours. If required for protein analysis, schizont material was harvested, washed in PBS, then harvested again and stored at -70°C until used. When returning labelled parasites to culture, schizont material was washed in RPMI w/Albumax.

In pulse-chase experiments, parasites were labelled as above for 10 minutes and returned to culture for 2 hours. Samples were taken at 0, 15, 30, 60 and 120 minutes following the return to culture and were washed in PBS, harvested, and stored at -70°C until used.

For labelling of GPI-anchored proteins, purified schizonts were resuspended to a 2-3% haematocrit in glucose-free RPMI w/Albumax supplemented with 10 mM fructose, 50 µCi ml^{-1} D-[6-^{3}H(N)]glucosamine hydrochloride and 50 µCi ml^{-1}

D-[2-^{3}H]mannose, and incubated at 37°C for 2 hours. Parasites were washed in PBS, harvested and stored at -70°C until used.

2.3.5 ^{35}S-labelling of parasites in the presence of brefeldin A

Magnet-purified schizonts were resuspended to a 2-3% haematocrit in RPMI w/Albumax in the presence of 5 μg ml^{-1} brefeldin A (BFA) from a stock solution of 1 mg ml^{-1} in methanol, or a corresponding amount of methanol alone for control cultures. The cultures were incubated for 1 hour at 37°C, then harvested and transferred to fresh medium, containing BFA or methanol but free of methionine and cysteine, except for a supplement of 100 μCi ml^{-1} Pro-Mix. Cultures were incubated for a further 1 hour at 37°C. The metabolically labelled parasites were washed in PBS, harvested by centrifugation and stored at -70°C until used.

2.3.6 Preparation of free merozoites

RPMI w/serum: As RPMI w/o, supplemented with 10% v/v pooled human AB+ serum (obtained from the National Blood Transfusion Service)

Tightly synchronised magnet-purified schizonts were returned to culture at a 2-3% haematocrit in RPMI w/serum. The culture was regularly examined by microscopy and merozoites were purified at an appropriate time. Most schizonts and uninfected erythrocytes were removed from the culture by centrifugation four times at 1600rpm/521*g* for 5 minutes. Supernatants containing merozoites and residual schizonts and erythrocytes were passed down the MACS depletion column as described above, and the flow-through containing merozoites was retained. Merozoites were harvested by centrifugation of the flow-through at 5000rpm/4670*g* for 5 minutes. Parasites were resuspended in PBS and transferred to microcentrifuge tubes, centrifuged again at 5000rpm/4670*g* and the merozoites stored at -70°C until used. When preparing merozoites for use in processing assays (cf. 2.7.7), the purification was carried out as above, except 2 mM EGTA was present at all stages following the elution of schizonts from the MACS column.

2.3.7 Preparation of culture supernatants

Tightly-synchronised magnet-purified schizonts were placed back into culture at a 2-4% haematocrit in RPMI w/o, and were closely monitored by thin blood smear until schizont rupture had occurred. Cellular material was pelleted and the supernatant clarified by successive 10 minute centrifugation steps at 2000rpm/814*g*, 5000rpm/4670*g* and 13,000rpm/16,060*g*. Supernatant was concentrated 10 x in a 20 ml centrifugal concentrator with a 5 kDa molecular weight cut-off, and stored at -20°C in 200 µl aliquots until used.

2.3.8 Cryopreservation and thawing of parasites

Freezing solution: 0.11 M NaCl, 0.17 M D-sorbitol, 28% v/v glycerol

High parasitemia ring-stage cultures were centrifuged at 1600rpm/521*g* for 8 minutes. Medium was aspirated and 300 µl aliquots of cells were transferred to CryoTubes (Nalgene). 700 µl freezing solution was added to each tube drop-wise, with gentle mixing. Tubes were immersed in $N_{2(l)}$ for storage.

To resuscitate, vials were taken from $N_{2(l)}$ and thawed at 37°C. The contents were transferred to a fresh tube. 0.2 ml of 12% w/v NaCl in PBS was added per ml of culture, drop-wise. 10 ml of 1.6% w/v NaCl in PBS per ml of culture was then added, followed by 10 ml 0.9% w/v NaCl, 0.2% w/v glucose in PBS per ml of culture - both solutions were added drop-wise. The culture was centrifuged at 1200rpm/293*g* for 5 minutes. The supernatant was aspirated and cells resuspended in 10 ml RPMI w/o. Following a further centrifugation at 1200rpm/293*g* for 5 minutes, the RPMI w/o was removed and replaced with RPMI w/Albumax, adding enough fresh erythrocytes to establish a 1-2% haematocrit culture.

2.3.9 Transfection of *P. falciparum*

Cytomix: 120 mM KCl, 0.15 mM $CaCl_2$, 2 mM EGTA, 5 mM $MgCl_2$, 10 mM K_2HPO_4/KH_2PO_4 (*p*H 7.6), 25 mM HEPES (*p*H 7.6)

WR99210: 20 µM stock (Jacobus Pharmaceuticals)

TE: 10 mM Tris-HCl, 1 mM EDTA (*p*H 8.0)

Sterile DNA for transfection was prepared by mixing 100 µg DNA, 1 ml ethanol and 40 µl 3 M sodium acetate (*p*H 5.2), and incubating at -20°C for 24 hours. The DNA mixture was then centrifuged at 13,000rpm/16,060*g* at 4°C for 30 minutes. The DNA pellet was washed with 500 µl cold, sterile 70% ethanol and centrifuged at 13,000rpm/16,060*g* at 4°C for 15 minutes. The DNA pellet was air-dried then resuspended in 30 µl sterile TE and stored at 4°C until required.

30 µl sterile DNA in TE was mixed with 370 µl sterile cytomix. DNA in cytomix was mixed with 200 µl packed parasitised erythrocytes (5-10% ring stage parasitemia) and transferred to a 0.2 cm electroporation cuvette. Parasites were pulsed at a voltage of 0.310 kV and capacitance of 950 µF, with resistance set to infinity, using a Bio-Rad Gene Pulser. Electroporated parasites were then transferred to a flask containing 10 ml RPMI w/Albumax at a haematocrit of 3%. Parasites were gassed and incubated at 37°C. WR99210 was added 24 hours after electroporation, to a final concentration of 20 nM. For the first 5 days post-transfection, both drug and medium were replaced daily.

Transfected parasites were cultured as normal, under drug pressure, until live parasites could be detected by thin blood smear. This point is referred to as the end of cycle 0. Parasites were then cultured for 3 weeks in the absence of drug pressure. Drug pressure was then re-applied until parasites re-emerged in the culture. This point is referred to as the end of cycle 1. This process of no drug pressure/drug pressure was repeated a further 2 times, until the end of cycle 3 had been reached. At the end of each cycle, parasite stocks were created (cf. 2.3.8) and genomic DNA prepared from the culture (cf. 2.4.1).

2.3.10 Cloning of mixed parasite population by limiting dilution

The parasitemia of a sorbitol-synchronised ring-stage culture was calculated by counting a thin blood smear. A haemocytometer was used to determine the number of cells, both infected and uninfected, present per ml of culture. The culture was diluted to a concentration of 3 x 10^{-3} parasites μl^{-1}, in RPMI w/Albumax with a 2% haematocrit, and 100 µl was aliquoted into the wells of a 96-well plate. This gave a theoretical number of 0.3 parasites per well. Plates were placed in a gassing chamber and gassed for 2 minutes, before incubation at 37°C.

Medium in the plates was first replaced after 5 days. This was repeated after 7 days, and again on day 12, with the addition of WR99210 at the usual concentration of 20 nM and an extra 0.5 µl packed erythrocytes per well. Medium was typically replaced every 5 days thereafter.

From day 21 onwards, thin smears were made from individual wells, and were examined for evidence of infection by Giemsa staining and light microscopy. If parasites were evident, the 100 µl cultures were transferred to a 6 well plate, in a volume of 5 ml RPMI w/Albumax and 2% haematocrit, and were ultimately cultured in 100 ml volumes.

2.4 Preparation and analysis of nucleic acids

2.4.1 Isolation of DNA from *P. falciparum*

Buffer A: 50 mM sodium acetate (*p*H 5.2), 100 mM NaCl, 1 mM EDTA

Parasitised erythrocytes were centrifuged at 1600rpm/521*g* for 8 minutes. The cell pellets were resuspended in 4 volumes Buffer A at room temperature, and mixed thoroughly. 1 volume 18% SDS (BioRad) was added and the cells mixed by flicking the tube, before a 2 minute incubation at room temperature. 6 volumes of a 25:24:1 mixture of phenol:chloroform:isoamyl alcohol were added to the samples, which were then mixed by inversion and centrifuged at 2000rpm/815*g* for 10 minutes. The aqueous layer was removed to a Corex tube and 1/10 volume 3 M sodium acetate (*p*H 5.2) and 2.5 volumes of cold ethanol were added. Samples were incubated overnight at -20°C.

DNA pellets were recovered by centrifugation at 10,000rpm/12,096*g* for 15 minutes at 4°C, and were air-dried following removal of supernatant. The pellets were resuspended in 500 µl TE, transferred to a 1.5 ml microcentrifuge tube, and 500 µl of a 25:24:1 mixture of phenol:chloroform:isoamyl alcohol was added. Samples were centrifuged at 8000rpm/6081*g* for 5 minutes. The upper aqueous layer was transferred to a fresh 1.5 ml tube and 500 µl cold isopropanol and 50 µl 3 M sodium acetate were added in order to precipitate the DNA. Centrifugation at 8000rpm/6081*g* for 10 minutes pelleted the DNA, which was then washed in 1 ml cold 70% ethanol, before a final spin at 8000rpm/6081*g* for 10 minutes.

Supernatants were discarded and the DNA pellet air-dried before resuspension in 50 µl TE and storage at 4°C in screw-top 1.5 ml tubes.

2.4.2 Agarose gel electrophoresis

DNA loading buffer: 25% v/v glycerol, 50% v/v 10 x TBE, 0.02% v/v bromophenol blue

DNA ladder: 1 kb or 100 bp DNA ladder (New England BioLabs) diluted in DNA loading buffer to 100 µg ml^{-1}. 0.5 µg loaded per lane.

Horizontal gels were made up by dissolving Agarose MP (Roche) in 0.5 x TBE with heating, typically to a concentration of 1% w/v. Ethidium bromide (Bio-Rad) was added to a concentration of 0.4 mg ml^{-1}. Gels were set in casts and combs (Pharmacia). DNA samples were mixed 4:1 with DNA loading buffer, loaded into gels and electrophoresed at a constant voltage of 4-5 V cm^{-1}, supplied by an EPS 500/400 DC Power Supply (Pharmacia). When sufficient separation had occurred, the DNA bands were visualised using a UV transilluminator. The DNA ladders described above allow an estimation of DNA concentration in samples. The mass of DNA in each band of the ladder is known, and by comparing the intensity of sample DNA bands with the intensity of a similar sized ladder marker, the mass of DNA in the sample loaded can be estimated.

2.4.3 Polymerase Chain Reaction

The *in vitro* amplification of specific DNA sequences was achieved through use of the polymerase chain reaction (PCR). Oligonucleotide primers were obtained from Sigma Genosys and the genomic DNA template was prepared as detailed in 2.4.1.

DNA fragments were typically amplified using AmpliTaq DNA polymerase (Applied Biosystems), or Expand (Roche) if a proof-reading enzyme was required, according to the manufacturers' instructions. A typical reaction volume of 50 µl consisted of 1 x PCR buffer, a final primer concentration of 0.4 µM, a final dNTP concentration of 0.2 mM, and 5 U AmpliTaq or 2.6 U Expand. The concentration of $MgSO_4$ varied between reactions but was typically 1.5 mM.

PCRs were carried out using an Eppendorf Mastercycler gradient machine, with programs similar to the following:

1. 94°C for 5 minutes
2. 94°C for 30 seconds
3. (T_m-5)°C for 30 seconds
4. 72°C for 1 minute per kb of sequence
5. Repeat steps 2-4 a further 29 times
6. 72°C for 10 minutes
7. Hold at 4°C

The annealing temperature of the primers (T_m) was estimated using the formula:
$T_m = [4 \times (G+C)] + [2 \times (A+T)]$
where G, C, A and T represent the number of bases of guanine, cytosine, adenine and thymine respectively, present in the oligonucleotide.

PCR products were separated and their size verified by agarose gel electrophoresis (cf. 2.5.2). When PCR products were to be used for cloning, they were purified from agarose using either the QIAquick or MinElute gel extraction kits (Qiagen), according to the manufacturer's instructions. QIAquick-purified DNA was eluted in 50 µl 10 mM Tris-HCl (*p*H 8.5) and MinElute-purified DNA in 10 µl 10 mM Tris-HCl (*p*H 8.5).

2.4.4 Restriction digest

DNA was digested with restriction endonucleases according to the manufacturer's protocol. All enzymes were obtained from New England Biolabs. DNA was typically incubated with enzymes in a total volume of 20 µl containing 1 x restriction buffer supplied by the manufacturer, supplemented with 100 µg ml^{-1} BSA if required. Restriction fragments were separated from unwanted DNA by agarose gel electrophoresis and the gel was visualised. If required, the band corresponding to the desired product was excised and purified from the gel slice using the QIAquick gel extraction kit or MinElute kit. Purified DNA was again electrophoresed alongside a molecular mass standard to allow an estimation of concentration to be made.

2.4.5 Southern blot analysis

1 x SSC:	150 mM NaCl, 300 mM $Na_3C_6H_5O_7$, *p*H 7.0
Denaturation buffer:	1.5 M NaCl, 0.5 M NaOH
Neutralising buffer:	1.5 M NaCl, 0.5 M Tris-HCl *p*H 7.4, 1 mM EDTA *p*H 8.0
Hybridisation buffer:	6 x SSC, 5 x Denhardts solution, 0.5% v/v SDS, 200 μg ml^{-1} herring sperm DNA
STE buffer:	100 mM NaCl, 10 mM Tris-HCl *p*H 8.0, 0.1 mM EDTA
Wash solution I:	6 x SSC, 0.1% SDS
Wash solution II:	2 x SSC, 0.1% SDS

Hybridisation buffer and wash solutions were heated to 65°C before use.

Genomic DNA was prepared from cloned parasite populations as described in 2.4.1. Equal quantities (~5 μg) of genomic DNA prepared from wild-type 3D7 parasites, parasites from cycles 0-3 and 5 parasite clones, namely 9, B4, F8, C7 and C9 were digested with *Sca*I and *Swa*I. Digested DNA was elecrophoresed on a large 0.8% agarose gel overnight at a constant voltage of 1 V cm^{-1}. The gel was shaken gently for 10 minutes in 0.25 M HCl, followed by a 5 minute wash in dH_2O. The gel was then shaken gently in denaturation buffer for 1 hour, with a change of buffer after 30 minutes, followed by 1 hour in neutralising buffer, with a change of buffer after 30 minutes.

Hybond-N membrane (GE Healthcare) was pre-wet in dH_2O then shaken for 15 minutes in 10 x SSC. A wick was made of two large sheets of Whatman 3MM paper on a solid support, the ends of the wick being in a tank of 10 x SSC. Four more layers of Whatmann 3MM paper, cut to the same size as the gel, were placed on top of the wick, and the gel was placed on top. The layer of Hybond-N membrane was placed on top of the gel and on top of the membrane was placed multiple layers of Whatman 3MM paper and paper towels. A weight held the set-up in place overnight to allow transfer of DNA to the membrane.

Following blotting, the membrane was shaken for 5 minutes in 0.4 M NaOH then for 1 minute in neutralising solution, after which the membrane was air-dried. DNA was crosslinked to the membrane by applying 1200 J using a Stratalinker UV crosslinker (Stratagene). The membrane was then wrapped in cling film and stored at -20°C until used.

The membrane was defrosted, washed in 2 x SSC, 0.1% SDS for 10 minutes and transferred to a hybridisation tube. The herring sperm DNA was heated to 95°C for 10 minutes and was added to 20 ml pre-heated hybridisation solution, to a final concentration of 200 μg ml^{-1}. Hybridisation solution was transferred to the hybridisation tube and the membrane was incubated at 65°C in a Hybaid oven for 3 hours.

Two probes were generated by two independent digestions of the transfection vector pHH1-PF08_0008HA3 (cf. 2.5.1). A double digest with *Xho*I and *Bgl*II generated an 841 bp probe, and a double digest with *Xmn*I and *Eco*RI generated an 807 bp probe. Following purification, probes were labelled using the DECAprime II Random Priming DNA Labeling Kit (Ambion) according to the manufacturer's instructions. DNA was denatured by boiling for 10 minutes in dH_2O before being added to the reaction mixture. Briefly, 30 ng denatured DNA was incubated with 1 x decamers, 1 x reaction buffer (-dATP), 5 U exonuclease-free Klenow enzyme, and 1.11 MBq [α-^{32}P]dATP (GE Healthcare) in a total volume of 25 μl, for 3 hours at 37°C. Labelled probes were then diluted to a volume of 50 μl with STE buffer and purified using ProbeQuant G-50 spin columns (GE Healthcare) according to the manufacturer's instructions. Following purification, probes were denatured by boiling for 5 minutes. The membrane was then incubated with 20 ml fresh hybridisation solution and a heat-denatured labelled probe, in a Hybaid oven overnight at 65°C.

Following removal of the labelled probe and hybridisation solution, membranes were washed at 65°C in the Hybaid oven with wash solution I for 30 minutes, twice. Two 30 minute washes with wash solution II were then carried out. The membrane was drained, wrapped in cling film and exposed to a sheet of Kodak Biomax MR film at -70°C for an appropriate length of time.

A plasmid integration event was confirmed by sequencing a PCR product amplified from the genomic DNA of a clonal population of parasites. The forward primer used was from 5' of the region of *PF08_0008* cloned into the pHH1-T996HA3 vector, and the reverse primer was from 3' of the triple HA tag coding sequence.

2.5 Construction of plasmids

2.5.1 Transfection construct

The *Xho*I- and *Bgl*II-digested pHH1-T996HA3 plasmid was a gift from Pippa Harris (NIMR Division of Parasitology) (Harris et al., 2005). *P. falciparum* 3D7 genomic DNA was used as a template to amplify an 829 bp DNA fragment from the 3' end of *PF08_0008* using oligonucleotide primers:
HAf (5'-GCGCAGATCTCCTACAGGTAA-3') and
HAr (5'-GCGCCTCGAGATTTAACAAGT-3').
Restriction sites introduced for cloning are underlined. *Xho*I and *Bgl*II sites were introduced to the 5' and 3' ends of the PCR product respectively. The fragment amplified consists of the extreme 3' 829 bp of the gene, excluding the terminal STOP codon. A $MgSO_4$ concentration of 4 mM was used for the amplification.

Gel-purified PCR product was digested with *Xho*I and *Bgl*II restriction endonucleases and ligated into the pre-digested pHH1-T996HA3 plasmid. The reaction volume of 10 µl consisted of 400 U T4 DNA ligase (New England Biolabs), 1 x T4 DNA ligase reaction buffer, digested vector and PCR product, with a vector:insert molar ratio of approximately 1:5. The ligation was carried out overnight at 4°C.

2.5.2 Ligation-independent cloning

Ligation-independent cloning (LIC) allows PCR products to be annealed into vectors without the need for restriction enzyme digest, through the use of vectors and PCR products with complementary overhangs. 2 regions of PF08_0008 were chosen for expression in a heterologous system using the LIC expression vectors pET 30 and pET 32 (Merck). One region is situated at the N-terminal end of the protein (N-terminal domain (NTD)) and one at the C-terminal end (C-terminal domain (CTD)). Regions of predicted low-complexity and hydrophobicity were avoided.

The pET 30 vector allows expression of recombinant protein as an N-terminal fusion with the His and S epitope tags. The pET 32 vector allows expression of recombinant protein as an N-terminal fusion with the His, S and thioredoxin epitope tags. The His and S tags enable the detection and purification of fusion

proteins, and the thioredoxin tag is believed to enhance the solubility of the fusion protein.

The coding sequence for the NTD and CTD regions was amplified from genomic DNA by PCR using the following sets of primers:
NTDf (5'-*GGTATTGAGGGTCGC*TGTGATATACAAAAAATAGCAG-3') and
NTDr (5'-*AGAGGAGAGTTAGAGCC*TTGTTCTTCATCTTCCTCAAT-3');
CTDf (5'-*GGTATTGAGGGTCGC*GCAAAAGCTTATTGTAAGAA-3') and
CTDr (5'-*AGAGGAGAGTTAGAGCC*GCCTTTGCATTTGGTCCTT-3').
The PCR was carried out as detailed in 2.4.3. The sequence in italics denotes LIC-compatible overhangs that are incorporated into the PCR product.

The PCR products corresponding to the N- and C-terminal regions of PF08_0008 were cloned into the pET 30 and 32 vectors respectively, according to the manufacturer's instructions. Briefly, 0.2 pmol of purified PCR product was treated with T4 DNA polymerase in the presence of dGTP to generate vector-compatible overhangs and allow efficient cloning into the LIC vector of choice through a simple annealing reaction.

2.5.3 Transformation of competent cells

S.O.C: 0.5% w/v yeast extract, 2.0% w/v tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM $MgSO_4$, 20 mM glucose.

Ligation mixtures were transformed into *E. coli* One Shot TOP10 competent cells (Invitrogen) according to the manufacturer's instructions. Briefly, 50 μl aliquots of cells were thawed on ice, and 1-4 μl ligation mixture was added. Cells were incubated on ice for 30 minutes and in a 42°C water bath for 30 seconds. Cells were returned to ice and 250 μl S.O.C. added. Following incubation at 37°C with shaking at 225 rpm for 1 hour, 200 μl of the transformation reaction was spread over a LB-Amp/Agar or LB-Kan/Agar plate as appropriate, pre-warmed to 37°C. pET 30 contains a gene encoding kanamycin resistance, and both pET 32 and pHH1-T996HA3 contain a gene encoding ampicillin resistance. Plates were inverted and incubated overnight.

2.5.4 Colony screening

A number of colonies were selected for a colony PCR screen. The conditions for the colony PCR were as initially used to amplify the insert. Reaction mixtures included the primers used to amplify the insert, however, no template was added to the reaction mixtures. Instead, individual bacterial colonies were picked using a disposable plastic colony picker. Cells were used to inoculate a referenced second LB Amp/Agar or LB Kan/Agar plate as appropriate, then dipped into tubes containing the PCR reaction mixture, cross-referenced with the second plate. PCR products from the selected colonies were analysed by agarose gel electrophoresis and positive colonies were identified by the presence of an appropriately-sized DNA fragment.

Plasmid DNA from colonies was isolated using a QIAprep Spin Miniprep kit (Qiagen), according to the manufacturer's instructions. Briefly, single colonies from the reference plate were used to inoculate 5 ml LB-Amp/Kan. Cultures were incubated for 12 hours at 37°C with shaking at 225 rpm. The cells from 1.5 ml of culture were pelleted by centrifugation at 13,000rpm/16,060*g* for 10 minutes, and the supernatant removed. The cell pellet was lysed under alkaline conditions. The cell lysate was neutralised and adjusted to high-salt conditions, which precipitated denatured proteins, cellular debris and genomic DNA. Centrifugation cleared the precipitate and the soluble plasmid DNA was then bound to the QIAprep silica membrane. Endonucleases and salts were removed in washing steps and DNA eluted in 50 µl 10 mM Tris-HCl, *p*H 8.5.

Restriction digests were carried out on the purified plasmid DNA samples using enzymes that cut either side of the ligated fragment. Digested plasmid DNA was subject to agarose gel electrophoresis, and the presence of an appropriately-sized fragment was indicative of a plasmid that had been successfully ligated to the insert.

Glycerol stocks of the positive colonies were made by mixing 200 µl glycerol with 800 µl of a 5 ml culture, inoculated from the reference plate as described above. Stocks were stored in CryoTubes at -70°C.

2.5.5 Sequencing of constructs and parasite genomic DNA

Sequencing of plasmid constructs was carried out using the Advanced Biotechnology Centre (ABC) DNA sequencing service at Imperial College, London. One μg plasmid DNA, or 120 ng PCR product, was supplied to ABC with 12.8 pmol of the appropriate primer in a final volume of 12 μl. Samples were sequenced following a BigDye v3.1 sequencing reaction and were subsequently run on an Applied Biosystems 3100 capillary sequencer.

In the case of the plasmid to be used for transfection, the sequencing primers used were:
CAMf 5'-TAAGAAAAACGAACATTAAGCTGCC-3' and
PbDTr 5'-GCATGCAAGCTTATAATTATTAATAGGTAC-3'.
These primers span the *XhoI*/*BglII* cloning site. Expression constructs were sequenced with the T7 promoter primer (Merck) on the sense strand, the primer
NTDsr 5'-TTGTTCTTCATCTTCCTCAAT-3'
for the antisense strand of the NTD pET 30 construct, and the primer
CTDr 5'-*AGAGGAGAGTTAGAGCC*GCCTTTGCATTTGGTCCTT-3'
for the antisense strand of the CTD pET 32 construct. PCR product amplified from parasites expressing PF08_0008-HA3 was sequenced with the forward primer
5'-CAGCTTTACAAGTTACAAAT-3'
and the reverse primer
5'-AAGAAAAACGAACATTAAGCTGCCATATC-3'.

Sequences were analysed using AutoAssembler software (Applied Biosystems).

2.5.6 Preparation of plasmid DNA for transfection

Plasmid DNA was isolated from a sequence-verified *E. coli* glycerol stock using the QIAfilter plasmid maxi kit (Qiagen). Cells were streaked onto a LB-Amp/Agar plate and incubated overnight at 37°C. Individual colonies were selected to inoculate 5 ml LB-Amp cultures, and cultures were incubated at 37°C, with shaking at 225 rpm, for 8 hours. 500 μl of these cultures were used to inoculate 500 ml LB-Amp cultures, which were incubated overnight under the same conditions.

Purification of plasmid DNA was carried out as per the kit protocol. Cultures were centrifuged at 4000rpm/4050*g* for 10 minutes to pellet the bacterial cells. The same alkaline lysis, neutralisation, clearing, binding and washing steps as the QIAprep Miniprep kit were carried out. DNA was eluted into a high-salt buffer then de-salted and concentrated by isopropanol precipitation. DNA pellets were washed in 70% ethanol, air-dried and dissolved in 500 µl dH_2O. Plasmid DNA was quantified by UV spectrophotometry at 260 nm.

2.6 Analysis of proteins

2.6.1 Total *Plasmodium falciparum* protein preparation

Reducing sample buffer:	1 M Tris-HCl *p*H 6.8, 4% v/v SDS, 20% v/v glycerol, 0.2 M DTT, 0.01% bromophenol blue
Non-reducing sample buffer:	1 M Tris-HCl *p*H 6.8, 4% v/v SDS, 20% v/v glycerol, 0.01% bromophenol blue

Purified schizont or merozoite material was resuspended in 2 x reducing or non-reducing sample buffer. Mixtures were passed through a 27 gauge needle, heated at 95°C for 5 minutes and stored at -20°C if not used immediately.

2.6.2 Extraction of membrane-associated proteins

High *p*H carbonate buffer: 0.1 M Na_2CO_3 *p*H 11.0

Purified schizonts or merozoites were resuspended in 20 volumes high *p*H carbonate buffer, passed through a fine-gauge needle and centrifuged at 50,000rpm/135,240*g* for 30 minutes at 4°C. Supernatants were removed and pellets were re-extracted with 20 volumes carbonate buffer followed by a further 30 minute spin at 50,0000rpm/135,240*g* at 4°C. Supernatants from both centrifugation steps were combined, buffer exchanged into PBS and concentrated by 50% using a centrifugal concentrator. Pellets and supernatant fractions were resuspended in 20 volumes 2 x reducing or non-reducing sample buffer and stored at -20°C if not used immediately.

2.6.3 Separation of proteins by SDS-PAGE

Coomassie:	0.1% w/v Coomassie blue, 45% v/v methanol, 10% v/v acetic acid
Destain:	45% v/v methanol, 10% v/v acetic acid
Resolving gel:	10% v/v acrylamide, 0.37 M Tris-HCl *p*H 8.8, 0.1% v/v SDS, 0.0026% v/v ammonium persulphate, 0.0005% v/v TEMED
Stacking gel:	3% v/v acrylamide, 0.12 M Tris-HCl, 0.1% v/v SDS, 0.05% v/v APS, 0.001% v/v TEMED
10 x Running buffer:	0.25 M Tris-HCl, 2 M glycine, 1% w/v SDS

Protein samples were separated by SDS-PAGE using the NuPAGE mini-gel system with the XCell Surelock Mini-Cell (Invitrogen). Pre-cast NuPAGE 10% or 12% Bis-Tris polyacrylamide gels (Invitrogen) were used with 1 x MES or 1 x MOPS running buffers as appropriate. Proteins to be separated were mixed with an equal volume of 2 x reducing or non-reducing sample buffer, and heated at 95°C for 5 minutes before loading. A protein standard (SeeBlue or SeeBlue Plus2 (Invitrogen)) was loaded for molecular mass estimation. A voltage of 200 V was applied until the dye front had reached the bottom of the gel, at which point it was removed from the gel system. Proteins could be visualised by staining the gel for 30 minutes with Coomassie blue with gentle shaking, and then by de-staining until complete.

For separation of immunoprecipitated proteins, large slab gels were used. Gels were typically 10% acrylamide. The resolving gel was poured into a vertical gel casting apparatus (Hoefer Scientific Instruments) to set for 1 hour, after which a stacking gel was made up and poured on top. A comb was inserted to create wells before setting. When set, gels were fixed to vertical electrophoresis apparatus and the reservoirs filled with 1 x running buffer. A constant current of 25 mA was applied until the dye front had reached the bottom of the gel. The gels were then fixed as detailed in 2.7.6.

2.6.4 Western blotting of proteins

Transfer buffer:	10% v/v methanol, 1 x NuPAGE transfer buffer
Washing solution:	0.1% v/v Tween-20 in PBS
Blocking solution:	10% w/v Marvel milk powder in washing solution
Antibody diluent solution:	5% w/v Marvel milk powder in washing solution

Proteins were separated by SDS-PAGE as described in 2.6.3. The blot was set up in an XCell II blot module (Invitrogen) containing the following components in order, all of which were pre-soaked in transfer buffer except for the gel: 2 gel-sized sponges, a piece of gel-sized Whatman 3MM paper, a piece of gel-sized Protran nitrocellulose membrane (Schleicher & Schuell), the SDS-polyacrylamide gel placed face down on the membrane, another piece of Whatman 3MM paper and 2 more sponges. The blot module was positioned vertically in an XCell SureLock Mini-Cell, with the blot set-up so that the nitrocellulose was on the same side of the gel as the anode. A voltage of 30 V was applied for 2 hours.

The membrane was removed and incubated with shaking at room temperature in blocking solution for 60 minutes. The membrane was then incubated with the primary antibody, diluted to the appropriate concentration, for one hour under the same conditions. The membrane was washed for 10 minutes in washing solution, three times. The nitrocellulose was then incubated for 30 minutes with a secondary antibody-HRP conjugate, diluted to the desired concentration. The membrane was then washed for 10 minutes, three times as before. Binding of antibody to the membrane was visualised by using enhanced chemiluminescence (ECL) western blotting detection reagents (GE Healthcare). Treated nitrocellulose was then exposed to Kodak Biomax MR film for an appropriate length of time and developed.

If required, the distance travelled along the gel by the bands corresponding to the molecular mass standards were measured and plotted against the $\log_{10}$(molecular mass). Standard curves could then be used to calculate the molecular mass of polypeptides using their gel migration distances.

2.6.5 Immunofluorescence assays (IFA)

Fixing solution:	4% v/v formaldehyde in PBS
Permeabilisation solution:	0.1% v/v Triton X-100 in PBS
Blocking solution:	3% w/v BSA in PBS
Antibody diluent:	1% w/v BSA, 0.05% v/v Tween 20 in PBS

Magnet-purified *P. falciparum* schizonts were used to make schizont smears on glass slides, which were wrapped in foil and stored at -20°C until required. Slides were dried at 37°C for 5 minutes and wells were marked using a water repellent marker pen. Slides were fixed for 30 minutes and washed for 5 minutes in PBS. Cells were then permeabilised for 10 minutes, washed in PBS for 5 minutes more and incubated overnight at 4°C in blocking solution.

Antibodies were diluted in the solution indicated above, and were used to probe slides for 60 minutes at 37°C under humid conditions. Dual labelling experiments were carried out by incubating slides with both antibodies simultaneously. Following incubation with primary antibodies slides were subjected to three five minute washes in PBS. Secondary antibodies were used to probe the slides for 30 minutes at 37°C under humid conditions. In the case of dual labelling experiments, secondary antibodies were incubated sequentially. Slides were subjected to three five minute washes in PBS and were briefly immersed in 0.5 mg ml^{-1} DAPI in PBS to stain nuclei. Slides were mounted in Citifluor, sealed and visualised under oil immersion using an Axioplan 2 Imaging system (Zeiss). Images were captured with Axiovision software and prepared for publication with Adobe Photoshop.

2.6.6 Immunoprecipitation of proteins

NP40 lysis buffer:	1% NP40, 50 mM Tris-HCl *p*H 8.0, 5 mM EDTA, 5 mM EGTA, 150 mM NaCl
SDS lysis buffer:	1% SDS, 50 mM Tris-HCl *p*H 8.0, 5 mM EDTA
Wash Buffer I:	1% NP40, 50 mM Tris-HCl, *p*H 8.0, 5 mM EDTA, 500 mM NaCl, 1 mg ml^{-1} BSA
Wash Buffer II:	1% NP40, 50 mM Tris-HCl, *p*H 8.0, 5 mM EDTA
Fixing solution:	25% v/v isopropanol, 65% v/v acetic acid in dH_2O

For the immunoprecipitation of protein complexes, pellets of metabolically labelled schizont material (~100 µl, prepared as detailed in 2.3.4 and 2.3.5) were resuspended in 1 ml NP40 lysis buffer supplemented with 1 x Complete protease inhibitors (Roche), and incubated on ice for 1 hour, followed by centrifugation at 13,000rpm/16,060*g* at 4°C for 25 minutes. Protein G Sepharose 4 Fastflow (GE Healthcare) was washed 4 times in NP40 lysis buffer and 150 µl 50% slurry was used to pre-absorb the soluble fraction of the parasite material by rotating for 1 hour at 4°C. Following centrifugation at 13,000rpm/16,060*g* for 2 minutes at 4°C, the supernatant was separated into 100 µl aliquots. 8 µl antiserum was added per aliquot, along with protease inhibitors to 1 x, and NP40 lysis buffer up to a total volume of 500 µl. Samples were rotated overnight at 4°C. For the immunoprecipitation of proteins from SDS-treated parasite material, the protocol was essentially the same, except schizont pellets were initially resuspended in 100 µl SDS lysis buffer and heated to 95°C for 5 minutes. NP40 lysis buffer was then added to 1 ml and the protocol continued as above, from the first centrifugation step.

50 µl 50% protein G slurry was added to samples the following day, which were then rotated at 4°C for 60 minutes to immunoprecipitate proteins. Samples were then centrifuged at 13,000rpm/16,060*g* at 4°C for 20 minutes, washed twice with wash buffer I and four times with wash buffer II. Following the removal of supernatant, samples were resuspended in 50 µl 2 x sample buffer and separated by SDS-PAGE. Gels for autoradiography were fixed for 20 minutes then shaken in Amplify (GE Healthcare) for 20 minutes, dried at 70°C for two hours using a slab gel dryer (Hoefer Scientific Instruments) and exposed to Biomax MR film at -70°C for an appropriate length of time.

2.6.7 Depletion of culture supernatant with erythrocytes

200 µl aliquots of 10 x culture supernatant were incubated in the presence or absence of 120 µl washed O+ erythrocytes, with shaking at 225 rpm at 37°C for 1 hour. Following centrifugation at 2000rpm/380*g* for 10 minutes, supernatants were removed, added to 120 µl fresh erythrocytes and incubated for a further hour. Samples were subjected to three successive centrifugation steps, at 2000rpm/380*g*, 8000rpm/6082*g*, and 13,000rpm/16,060*g*, each for 10 minutes, with the supernatant being retained at each step and cellular material discarded. 2 x sample buffer was added to samples, which were then heated to 95°C for 5 minutes, separated by SDS-PAGE and subject to Western blot analysis.

For the elution of proteins bound to the surface of erythrocytes, 400 µl aliquots of supernatant were incubated with 60 µl O+ erythrocytes for 1 hour at 37°C. The supernatant was removed following centrifugation at 2000rpm/380*g* for 10 minutes, replaced with 400 µl fresh supernatant and the incubation repeated. Erythrocytes were isolated through centrifugation at 2000rpm/380*g* for 10 minutes, and were washed three times in PBS. 20 µl 2 M NaCl was added to the cells drop-wise to elute proteins from the erythrocyte surface. Following a further centrifugation step the supernatant was recovered and resuspended in 2 x reducing sample buffer. Eluted proteins were separated by SDS-PAGE and subject to Western blot analysis.

2.6.8 Merozoite processing assay

Processing buffer: 50 mM Tris-HCl *p*H 7.6, 5 mM $CaCl_2$

Processing of proteins on the surface of merozoites can occur independently of erythrocyte invasion. The preparation of merozoites in the presence of the calcium-chelating agent EGTA reversibly inhibits this processing.

Merozoite pellets, prepared in the presence of EGTA, were removed from storage at -70°C and immediately placed into ice. The cells were washed three times in 1 ml ice cold processing buffer, with centrifugation at 13,000rpm/16,060*g* for 2 minutes at 4°C. Pellets were resuspended in 150 µl processing buffer and divided into aliquots of 18 µl in pre-chilled microcentrifuge tubes on ice. 2 µl processing buffer or 100 mM EGTA were added separately to frozen aliquots, which were

thawed and immediately returned to ice for 5 minutes. Merozoites were then incubated at 37°C for 2 hours. After the 2 hour incubation another aliquot was taken from -70°C to ice, and 2 µl 0.5 M EGTA was added – this was the control sample (t=0) which would demonstrate the background level of processing that had already occurred in merozoite samples before incubation at 37°C. The control merozoite sample was thawed as above and all samples were centrifuged at 13,000rpm/16,060*g* for 5 minutes at 4°C. 15 µl supernatant was removed from each sample and mixed with an equal volume of 2 x reducing sample buffer, boiled for 5 minutes then separated by SDS-PAGE and subject to Western blot analysis. Densitometry was performed on the resultant bands using the ImageJ program (Wayne Rasband, National Institutes of Health), in order to estimate the level of processing occurring in samples relative to the t=0 control.

2.7 Expression of recombinant PF08_0008 fragments

2.7.1 Expression host transformation

When colonies containing the correct-sized insert had been identified, their plasmid DNA was transformed into an expression host, *E. coli* BL21 (DE3) cells (Merck), according to manufacturer's instructions. Briefly, cells were thawed on ice, 20 µl were transferred to pre-chilled polypropylene microcentrifuge tubes and 1 µl plasmid DNA was added. Cells were incubated on ice for 5 minutes, then for 30 seconds in a 42°C water bath. Cells were returned to ice for 2 minutes then 80 µl S.O.C was added. Following incubation at 37°C with shaking at 225 rpm for 1 hour, 50 µl of the transformation was spread over a pre-warmed LB Amp/Agar or LB Kan/Agar plate as appropriate. Plates were incubated overnight.

2.7.2 Small-scale recombinant protein solubility test

5 ml LB-Kan or LB-Amp as appropriate were inoculated with a colony from the expression host plates and incubated overnight at 37°C with shaking at 225 rpm. 50 µl of the overnight culture was added to 1 ml LB-Kan/Amp and the cultures incubated in microcentrifuge tubes for 1 hour at 37°C, with shaking at 225 rpm. Cultures were split into two 500 µl aliquots and IPTG added to a final concentration of 1 mM to one of the aliquots. All cultures were incubated for a further 90 minutes at 37°C, then centrifuged at 13,000rpm/16,060*g* for 10 minutes

to pellet the cells. Culture supernatants were removed and the pellet resuspended in 300 µl BugBuster (Merck), with the addition of 1 µl Benzonase Nuclease (Merck). Cell lysates were shaken at room temperature for 20 minutes then centrifuged at 4°C for 20 minutes at 13,000rpm/16,060*g*. The soluble protein fraction was removed from the pellet into a separate tube. Pellet fractions were resuspended in 300 µl 27% v/v BugBuster in dH_2O and all fractions were stored at -20°C until used.

Proteins were separated by SDS-PAGE and visualised by Coomassie staining as described in 2.6.3. Western blots were carried out as detailed in 2.6.4 using a rabbit anti-His primary antibody (Santa Cruz Biotechnology) and a goat anti-rabbit HRP-conjugated secondary antibody (Bio-Rad). Bacterial clones that were shown to express a correctly sized and tagged recombinant protein were stored for future use as glycerol stocks (cf. 2.5.4).

2.7.3 Large-scale expression and purification of recombinant protein

Buffer B: 8 M urea, 0.1 M NaH_2PO_4, 10 mM Tris-HCl, *p*H 8.0
Washing buffer: 50 mM NaH_2PO_4, 300 mM NaCl

For large scale purification of insoluble protein, glycerol stocks were streaked out onto LB-Kan/Agar or LB-Amp/Agar plates as appropriate and incubated overnight. 5 ml overnight cultures were generated from individual colonies on the resultant plates, and used to inoculate 1 litre cultures of LB-Kan/LB-Amp, which were incubated at 37°C with shaking until an OD of 0.6 was reached. Cultures were then induced with 1 mM IPTG and grown for 4 hours.

Cultures were harvested by centrifugation at 4000rpm/4050*g* to pellet the cells. Cells were lysed in BugBuster in the presence of Benzonase and 1 x Complete protease inhibitors, according to manufacturer's instructions. The resultant soluble protein fraction was discarded. The pellet of insoluble protein was resuspended in 50 ml Buffer B and incubated at room temperature with shaking until the suspension had cleared. The suspension was centrifuged at 13,000rpm/20,442*g* for 20 minutes at 4°C to separate residual insoluble protein, and the supernatant was saved for purification.

Disposable plastic gravity-flow columns (Pierce) were packed with Ni-NTA agarose (Qiagen) and equilibrated with 3 column volumes of Buffer B. The supernatant was passed down the column and flow-through collected. The column was subsequently washed with 10 column volumes of washing buffer supplemented with 10 mM imidazole, 5 column volumes of washing buffer containing 20 mM imidazole and finally 5 column volumes of washing buffer containing 50 mM imidazole. The His-tagged protein was eluted into 3 column volumes of washing buffer with 250 mM imidazole. All wash steps were collected and were analysed by SDS-PAGE and Coomassie staining, along with the flow-through and eluted fractions.

When the purity of the eluted fraction had been confirmed by SDS-PAGE, purified protein was dialysed against 3 litres of PBS, with 2 changes, using SnakeSkin dialysis tubing (Pierce) over a period of approximately 24 hours. Equal volumes of dialysed protein and BSA protein standards of known concentrations were analysed by SDS-PAGE. Following Coomassie staining, the relative intensities of the bands corresponding to the protein standards were calculated by densitometry using the ImageJ program, and used to calculate the concentration of the two recombinant proteins.

2.8 Generation of polyclonal antisera

For each antigen, five eight-week old female Balb/c mice per antigen were immunised intra-peritoneally with 50 μg protein mixed 1:1 with Freund's complete adjuvant. Mice were boosted 3 times at 3 week intervals with 100 μg protein in Freund's incomplete adjuvant. 3 weeks after the final boost, mice were killed according to a schedule 1 method. Serum was obtained from total blood by centrifugation at 6000rpm/3421*g* for 20 minutes at 4°C, then stored at -20°C. Mice immunisations were carried out with the help of Sola Ogun (NIMR, Division of Parasitology).

New Zealand White rabbits were immunised subcutaneously by Harlan Sera Lab in accordance with their 77 day immunisation protocol. 3 rabbits per antigen were immunised with 200 μg protein in Freund's complete adjuvant, then boosted 5 times at 2 week intervals with 200 μg protein in Freund's incomplete adjuvant. Antibodies were obtained from Harlan as IgG, purified from the terminal bleeds

with protein A, and were stored in the presence of 0.1% sodium azide at -20°C. Sodium azide-free antibodies were obtained by buffer exchange into PBS using a PD-10 desalting column (GE Healthcare) according to the manufacturer's instructions, followed by a centrifugal concentrator.

2.8.1 Purification of IgG

1 ml HiTrap Protein G columns (GE Healthcare) were used according to the manufacturer's instructions, to purify IgG from antisera and pre-immune sera if necessary. Briefly, columns were washed with 10 ml PBS, the serum was applied and the column washed with a further 10 ml PBS. IgG was eluted with 5 ml 0.1 M glycine, *p*H 2.7, in 500 µl fractions, into tubes containing 100 µl 1.0 M Tris-HCl *p*H 9.0. Samples of each fraction were analysed by SDS-PAGE and Coomassie staining, and fractions containing IgG were pooled and buffer exchanged into PBS.

2.9 Invasion inhibition assays

Parasites were synchronised using a two-stage magnet purification as detailed in 2.3.3. Parasites were cultured in type O+ human erythrocytes for invasion assays. Following synchronisation, cultures were diluted in fresh blood to a parasitemia of 0.5%, with a 1% haematocrit. At approximately 36 hours post-invasion, 90 µl parasite culture were seeded into 96-well plates, in the presence of PBS, 1 mg ml^{-1} pre-immune rabbit IgG, 12.5 mM EGTA, or either of the test rabbit antibodies anti-CTD or anti-NTD, with a total culture volume of 100 µl. Antibodies were diluted in PBS. 100 µl of a 1% haematocrit culture of uninfected erythrocytes was also plated. Parasite cultures were seeded in triplicate. Test antibodies were used at final concentrations of 1 mg ml^{-1}, 0.5 mg ml^{-1} and 0.25 mg ml^{-1}, and were diluted in rabbit pre-immune IgG to a final concentration of 1 mg ml^{-1} total IgG as appropriate.

Following incubation at 37°C for approximately 42 hours in a gassed, humidified chamber, 50 µl of each culture was resuspended in 200 µl 10 µg ml^{-1} ethidium bromide in PBS, and cells were stained in the dark for one hour. Cells were pelleted by centrifugation at 4000rpm/1520*g* for one minute, the supernatant was removed and cells were resuspended in 400 µl PBS. Samples were processed using a FACSCalibur flow cytometer (BD Biosciences). Parasitemia was

calculated using the CellQuest Pro software (BD Biosciences) by first gating for intact erythrocytes by side scatter and forward scatter, and then determining the proportion of ethidium bromide-positive cells. Mean parasitemia and standard deviation were calculated and used to determine the invasion efficiency for each condition relative to the PBS control.

3. *Identification of PF08_0008 as a gene for study*

3.1 Identification of a candidate gene for study from proteomic data

Data from the Florens *et al.* proteomic study of 2002 were analysed in order to begin the identification of a candidate gene for study. This study confidently identified 2415 *P. falciparum* proteins across four stages of the parasite life cycle (sporozoite, merozoite, trophozoite and gametocyte), through multi-dimensional protein identification technology (MudPIT).

The raw data from this study were filtered according to a set of criteria designed to identify those proteins likely to play a role in the invasion of host cells, and this is illustrated in figure 3.1a. In the first two steps, proteins possessing a signal peptide and a single transmembrane domain were 'pulled out' of the data set. No additional analysis of the raw data was required at this stage, as signal peptide and transmembrane domain predictions were already present in the data set. This stage of the filtering process identified 92 proteins and should include proteins that are targeted either to the secretory organelles or to the parasite surface, as is typical of proteins known to be involved in invasion. The presence of a single transmembrane domain is also typical of proteins involved in invasion such as AMA1, and the PfRH and DBL-EBP protein families. AMA1 and EBA140 were included in this set of 92 proteins, as expected given that both proteins possess a signal peptide and a single transmembrane domain - this indicated that the initial stages of the filtering process were working as expected.

A broad functional classification of each identified protein was included in the original dataset, and of the 92 proteins that were predicted to possess a signal peptide and single transmembrane domain, only 40 were assigned as hypothetical. This set of 40 proteins was filtered further to identify those that were detected in merozoite and sporozoite protein preparations only. The rationale for including this step was that the basic machinery of invasion is believed to be conserved between merozoites and sporozoites (Baum et al., 2006).

From the 2415 proteins initially identified in this proteomic study, only two were identified that fulfilled all of the set criteria. These were PF08_0008 and PF14_0016.

***Figure 3.1*: Identification of an uncharacterised gene for study from proteomic data**

a) Data from the Florens *et al.* proteomic study of 2002 were filtered according to a set of criteria. The number of proteins initially identified in the study was 2415. This number was reduced by identifying only those that fulfilled a set of 4 criteria. The characteristics required of a protein for it to pass through each stage of data filtering are shown on the right of the figure. The presence of a transmembrane domain and signal peptide were predicted using TMHMM (Krogh et al., 2001) and SignalP (Nielsen et al., 1997), respectively. The number of proteins fulfilling each criteria are shown in the coloured boxes. The white boxes contain the PlasmoDB Gene ID of the two proteins that fulfil all of the criteria listed.

AMA1 and EBA140, which both contain a single transmembrane domain and possess a signal peptide, were present after the second filter was applied, in the set of 92 proteins. They were not present in the dataset after the third filter was applied, which only allowed hypothetical proteins to enter the next stage, thus verifying the initial stages of the filtering process.

b) The amino acid sequence of PF08_0008. The signal peptide predicted by SignalP (residues 1-21) is indicated in red, and the transmembrane domain predicted by TMHMM (residues 715-737) is indicated in blue.

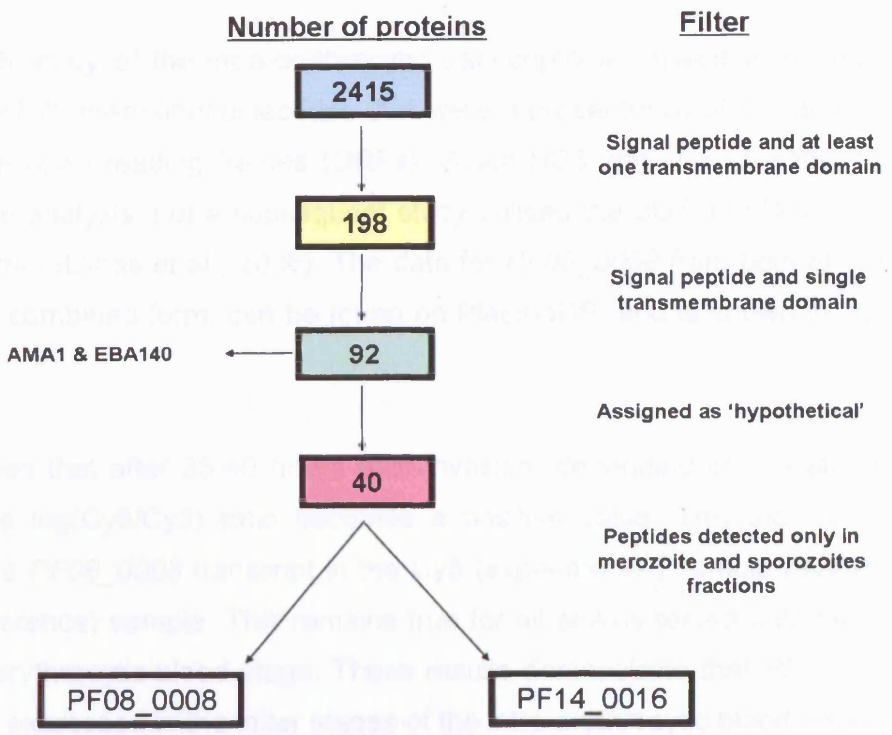
Further investigation into these two candidates was carried out. The PF14_0016 and PF08_0008 protein sequences were retrieved from PlasmoDB, and a BLAST search was carried out to determine if either of these protein sequences shared any similarity with other proteins, whether from *Plasmodium* or other species.

When a combined BLAST and conserved domain search is carried out on the amino acid sequence of PF08_0008 using the NCBI-BLASTP server, the top four significant hits are with hypothetical proteins from *P. vivax*, *P. chabaudi*, *P. berghei*, and *P. yoelii*. There are no significant hits with proteins from other species. This result suggests that PF08_0008 is conserved across different *Plasmodium* species, and could imply a conserved function for the protein in the parasite life cycle. No conserved domains were detected in PF08_0008.

Three of the top four BLAST hits for PF14_0016 were with members of the early transcribed membrane protein (ETRAMP) family (Spielmann & Beck, 2000). Upon further investigation it appears that PF14_0016 is a member of this protein family and is encoded by *etramp*14.1 (Spielmann et al., 2003). Importantly, ETRAMPs are small, highly charged integral parasitophorous vacuole membrane (PVM) proteins, that are believed to be important in the formation of different microdomains at the parasite-erythrocyte interface, post-invasion (Spielmann et al., 2006). These proteins are therefore of no direct relevance in the search to find a protein involved in invasion.

The focus of this study has therefore become the protein encoded by *PF08_0008*. The amino acid sequence of PF08_0008 is shown in figure 3.1b.

a)



b)

```
MKYYTSLYVALIIALCQAVSALIRNSNTPQAFLIPELNNNEKNEFNNNEKNEMNNNLNNE 60
FNNNEENCDIQKIAEEMMENLLNENDMYTNIMLSLQNRLSDDYLCSEPKYENICIHEKDK 120
ISLSFPCSNPKYEKLIHKFTFKKLCNSKAAFNNTLLKSFIEEDEEQNTFSLMLKQFKILL 180
TCVDDELKDIYKESIDLLVDLKTSITELTQKLWSGKMVNVLKKREFLITGILCELRNGNK 240
SPLISKSLEFENLGILKMNNEELLNEAYNAFSDYYYFFPYFIQKLLEKGGMIERLIKIHE 300
NLTKYRTKDMVNKINAQSKGEVLNNEDILNKLNAYKHYTKHGATSFIQSREVKIVNQNVN 360
NDDTTKNQQQNVNNNEKLNNNNNNNNNQQVNNNNNNNNNQQVNNNNNNNNNQVNNNNNNN 420
NQVNNNNYNNNNQVNNNNNNNQQVNNNNNYNNQLNNNNFNNNLQVNKNDKHVPKNNHTTA 480
THTNNILYNPLYSINPEKPKDIIKLLKDLIKYLHIVKFENNEPTTNIDEEGIRKLLENSF 540
FDLNDDILIVRLLLKPQTVILTVIQSFMLMTPSPSRDAKAYCKKALINDQLVPTNDTNIL 600
SEENELVNNFSTKYVLIYEKMKLQELKEMEESKLKMKYSKTNLSALQVTNPQNNKDKNDA 660
SNKNNNPNNSSTPLIAVVTDLSGEKTEDIINNNVDIATLSVGVQNTFQGPNAKAGSLINH 720
LSYATFLFFSFILINLLN
```

3.2 Transcriptome data

In order to provide additional support to the case for choosing PF08_0008 as the protein for study, information was retrieved from the two genome-wide transcription studies that were published in 2003 (Bozdech et al., 2003, Le Roch et al., 2003).

The Bozdech study of the intra-erythrocytic transcriptome utilised a microarray with probes of 70-mer oligonucleotides that were representative of the predicted *P. falciparum* open reading frames (ORFs). Strain HB3 was used for the initial transcriptome analysis, but a subsequent study utilised the 3D7 and Dd2 strains of *P. falciparum* (Llinas et al., 2006). The data for *PF08_0008* from both of these studies, in a combined form, can be found on PlasmoDB, and is shown in figure 3.2a.

It can be seen that after 35-40 hours post invasion, depending on the strain in question, the log(Cy5/Cy3) ratio becomes a positive value. This indicates an excess of the *PF08_0008* transcript in the Cy5 (experimental) sample relative to the Cy3 (reference) sample. This remains true for all strains tested until the end of the intra-erythrocytic blood stage. These results demonstrate that *PF08_0008* is maximally expressed in the latter stages of the intra-erythrocytic blood cycle.

The authors of the initial transcriptome study (Bozdech et al., 2003) used the data generated to predict a set of proteins that may be important as vaccine candidates or in the process of invasion. The expression profiles of seven of the best current malaria vaccine candidates (AMA1, MSP1, MSP3, MSP5, EBA175, RAP1 and RESA1), some of which are known to play a role in invasion, were compared with the entire transcriptome. ORFs were ranked based on the similarity of their expression profile to those of the seven vaccine candidates and the top 5% (262 ORFs) were highlighted. This gene set contained 28 genes previously associated with antigenicity or merozoite invasion. However, 189 genes in this set are currently of unknown function, and *PF08_0008* is one such gene.

The Le Roch study utilised a custom oligonucleotide array including 260,596 25-mer probes from predicted coding sequences, with probes placed, on average, every 150 bases on both DNA strands across the entire *P. falciparum* genome. In

addition to analysing asexual intra-erythrocytic forms of *P. falciparum*, this study also used RNA prepared from sporozoites, merozoites and gametocytes.

Data for *PF08_0008* from this study is shown in figure 3.2b. Absolute expression values published were calculated using the match-only integral distribution (MOID) algorithm. Since the chip used contained multiple probes per gene, a quantitative estimation of a gene's transcript abundance can be obtained and used to compare expression between stages. For this array, moderate levels of gene expression are indicated by absolute values in the region of 100-1000, with low levels of expression indicated by values less than 100 and high expression levels by values above 1000.

Results from both the temperature- and sorbitol-synchronised methods appear to show that the maximal expression of *PF08_0008* occurs in the schizont and sporozoite stages. Merozoites prepared from sorbitol-synchronised parasites also showed high levels of *PF08_0008* expression relative to other stages. Again, this would be typical of a gene that is involved in invasion. Using the above classification, the level of gene expression for *PF08_0008* is moderate.

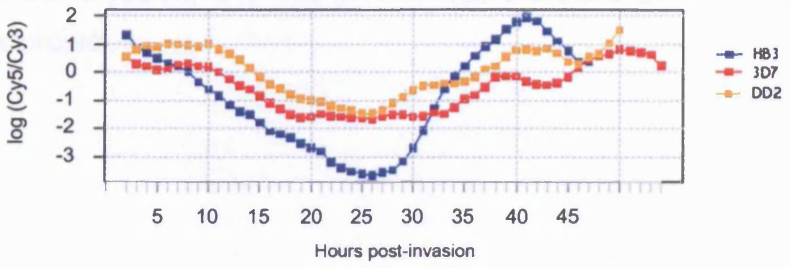
One of the aims of the Le Roch study was to demonstrate that genes with similar function have similar expression profiles. This was with a view to rapidly assigning functions, with a reasonable level of confidence, to the many thousands of uncharacterised genes. Cluster analysis was performed on the data based on the timing of gene expression, and it was found that membership in a particular cluster was highly predictive of the cellular function of the gene product. For example, cluster 15 from this study contained 130 genes that were highly transcribed at the schizont stage and contained most genes that have known or presumed roles in cell invasion including *EBA140*, *PfSUB2*, six of the eight merozoite surface proteins (MSPs) and *AMA1*. Importantly, *PF08_0008* was also placed in this cluster.

Figure 3.2*: Published transcriptome data available for *PF08_0008

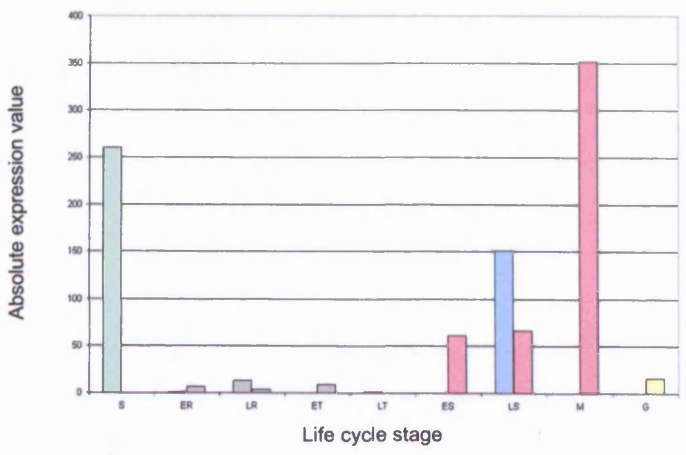
a) The expression profile of *PF08_0008* throughout the intra-erythrocytic blood stage is shown for *P. falciparum* strains 3D7 (red), HB3 (blue) and Dd2 (orange), the result of two spotted microarray studies (Bozdech et al., 2003, Llinas et al., 2006). Log ratios of Cy5 (experimental time-point cDNA) to Cy3 (reference pool of cDNA from all time-points) are shown on the y-axis. This figure is a modified form to that found at the PF08_0008 entry at http://plasmodb.org.

b) Graph showing the absolute expression values of *PF08_0008* for different lifecycle stages of *P. falciparum* (Le Roch et al., 2003). Expression values were calculated using the match-only integral distribution (MOID) algorithm. Results for asexual blood stages synchronised using two different methods are shown in different colours - sorbitol synchronisation results in pink and temperature synchronisation results in blue. Result for sporozoites shown in green. Grey values indicate results that were below the confidence threshold. S = sporozoites, ER = early rings, LR = late rings, ET = early trophozoites, LT = late trophozoites, ES = early schizonts, LS = late schizonts, M = merozoites, G = gametocytes.

a)



b)



3.3 Conclusions

The dataset of the Florens *et al.* (2002) proteomic study was filtered to identify uncharacterised *P. falciparum* proteins, detected only in merozoite and sporozoite fractions, that possess a predicted signal peptide and single transmembrane domain and thus have characteristics of *P. falciparum* proteins known to be involved in invasion. Of 2415 proteins present in the initial dataset, only two fulfilled all of the criteria: PF08_0008 and PF14_0016, the latter being ultimately excluded since it is a member of the ETRAMP family of proteins that function post-invasion.

Data from both of the transcriptional studies described above concur, and suggest a late blood-stage expression profile for *PF08_0008*. Both sets of transcriptome data also support the result from the proteomic study indicating PF08_0008 is present in merozoites, and the Le Roch study supports the proteomic result indicating the presence of the protein in sporozoites.

Taken together, the available proteome and transcriptome data show that PF08_0008 displays the characteristics expected of a protein involved in the process of host cell invasion, and this protein was therefore the focus of the remainder of the project.

4. Bioinformatic characterisation of PF08_0008

4.1 Introduction

Following the choice of gene for study, a variety of bioinformatics tools were used in order to gather information on PF08_0008 that could be important in the design of future experiments. Published high-throughput studies that have been used to identify *P. falciparum* GPI-anchored proteins and protein-protein interactions are also discussed below.

4.2 PF08_0008 is not predicted to be targeted to the apicoplast

The PlasmoAP tool, available at http://plasmoDB.org was used in order to predict the likelihood that PF08_0008 is targeted to the relict, non-photosynthetic plastid in the parasite that is known as the apicoplast. Transport to the apicoplast is dependent on the presence of a bipartite signal, consisting of a classical N-terminal signal sequence, followed by a transit peptide. The transit peptide of apicoplast-targeted proteins has certain characteristics such as a net positive charge and an enrichment of lysine and asparagine residues. These rules were used to develop the PlasmoAP tool (Foth et al., 2003). When the PF08_0008 amino acid sequence is input, the server predicts that it is unlikely to be an apicoplast-targeted protein.

4.3 Predictions of transmembrane domains within PF08_0008

Four web-based algorithms were used to determine whether PF08_0008 is likely to possess any transmembrane domains. The results are summarised in figure 4.1a. There appears to be no overall consensus between the programs used, although the same two hydrophobic regions of the protein are consistently recognised as potential transmembrane regions. These regions correspond to two of the three most prominent peaks on the hydrophobicity plot of PF08_0008 (figure 4.1b).

All four algorithms also predicted a transmembrane domain at the extreme N-terminus of the protein. However, the SignalP server (http://www.cbs.dtu.dk/services/SignalP/) predicts with a probability of 0.985 that

the initial 21 residues of the PF08_0008 sequence represent a classical signal peptide for entry into the secretory pathway, with cleavage between residues 21 and 22. Consequently these extreme N-terminal transmembrane predictions are omitted from figure 4.1a.

As detailed in chapter 1, transmembrane domains consist of a hydrophobic region followed by a 'stop transfer' signal consisting of charged residues. The sequence of the transmembrane region predicted by TMAP, TOPPRED2 and TMPRED, between residues 550 and 578, is shown in figure 4.1c i. 16 of the 25 residues between positions 550 and 574 are hydrophobic, and there are also residues that could serve as a 'stop transfer' signal located just downstream of the predicted transmembrane region. This transmembrane region is typical in location and size of those found in proteins such as EBA175, AMA1 and PTRAMP, molecules thought to be involved in invasion. However, if a true type I membrane protein, the cytosolic tail of PF08_0008 would be slightly longer than is typical of these other type I proteins (figure 4.1d).

It is unlikely that the extreme C-terminal hydrophobic region recognised by TMHMM2, TOPPRED2 and TMPRED represents a true classical transmembrane domain. As can be seen from figure 4.1c ii, the hydrophobic region extends almost to the C-terminus. This would mean that the cytosolic domain of the protein would be completely lacking, or would consist of a single asparagine residue, and this is unlikely to provide a sufficient 'stop transfer' signal or to be of any functional or structural relevance. As mentioned in chapter 1, it is possible that a predicted transmembrane domain at the extreme C-terminus of a protein is a mis-assigned GPI-anchor addition sequence, and this is discussed further below.

TOPPRED2 and TMPRED both predicted two transmembrane domains. For both algorithms, the prediction of a transmembrane domain at the extreme C-terminus received a higher score than the prediction that placed a transmembrane domain in the residues 555-575 region.

***Figure 4.1*: Bioinformatic predictions of transmembrane domains within PF08_0008**

a) Four web-based algorithms were used to predict the location of transmembrane domains, if any, within the amino acid sequence of PF08_0008. A schematic view of PF08_0008 is shown for each algortihm, along with the location and size of the transmembrane domain(s) they predicted indicated in pink. The residues predicted to be part of a transmembrane domain are detailed below the name of the algorithm on the right of the figure.

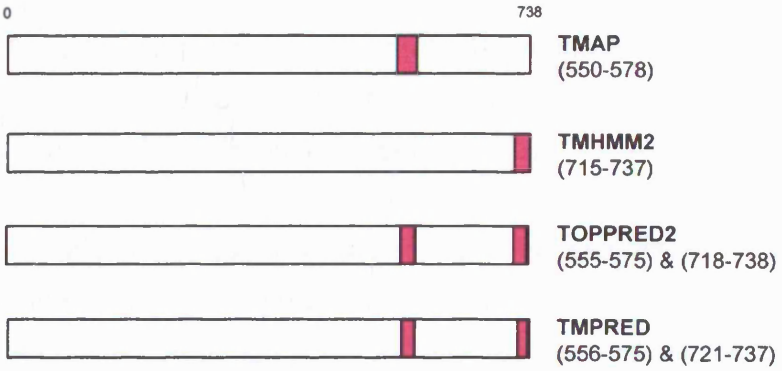
b) A hydrophobicity plot was generated for PF08_0008, based on the Kyte & Doolittle (1982) method, using the biological sequence alignment editor BioEdit. Mean hydrophobicity for a window size of 13 is shown on the y-axis, and amino acid position is shown on the x-axis. The two asterisks mark the two most prominent peaks, excluding the peak at the N-terminus that corresponds to the predicted signal peptide.

c) i) The amino acid sequence of PF08_0008 from residues 550 to 587 is shown. The transmembrane domains predicted by the TMAP, TOPPRED2 and TMPRED algorithms are contained within this region. Hydrophobic residues are indicated in red. Charged residues that follow the hydrophobic region, and could act as a 'stop transfer' sequence, are underlined.

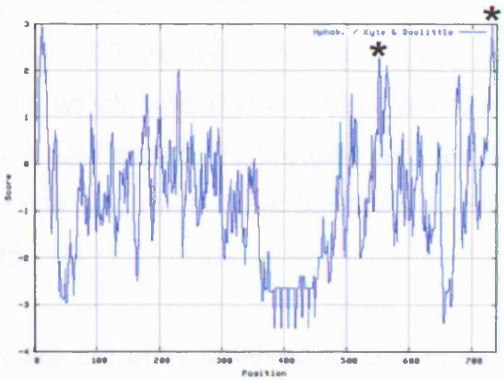
ii) The amino acid sequence of PF08_0008 from residues 710 to the C-terminus (position 738) is show. The transmembrane domains predicted by the TMHMM2, TOPPRED2 and TMPRED algorithms are contained within this region. Hydrophobic residues are indicated in red.

d) A comparison of the size and location of the transmembrane domains (shown in pink) of PTRAMP, AMA1 and EBA175. PF08_0008 is also shown with one of the transmembrane domains predicted by TOPPRED2 (residues 555-575).

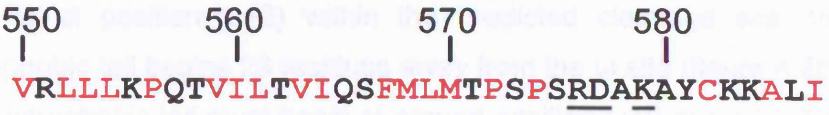
a)


0
738
TMAP
(550-578)
TMHMM2
(715-737)
TOPPRED2
(555-575) & (718-738)
TMPRED
(556-575) & (721-737)

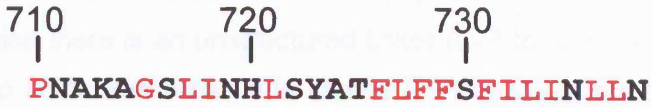
b)


*
*

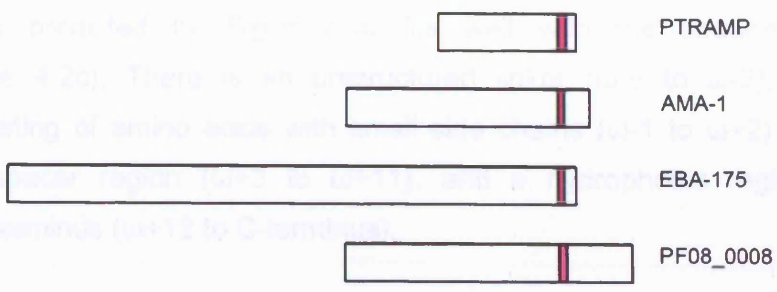
c) i)


550
560
570
580
VRLLLKPQTVILTVIQSFMLMTPSPSRDAKAYCKKALI

ii)


710
720
730
PNAKAGSLINHLSYATFLFFSFILINLLN

d)


PTRAMP
AMA-1
EBA-175
PF08_0008

4.4 Predictions of GPI-anchorage of PF08_0008

Two web-based GPI prediction servers were used to determine the likelihood of PF08_0008 being a GPI-anchored protein: the Big-π GPI prediction server (Eisenhaber et al., 1999) and DGPI (Kronegg & Buloz, 1999).

The results are summarised in figures 4.2b and 4.2c. No sequences scored highly enough for the Big-π GPI prediction server to confidently assign PF08_0008 as GPI-anchored, although it did identify residue 714 as the best possible cleavage (ω) site.

DGPI predicted GPI-anchorage of PF08_0008 with 2 potential cleavage (ω) sites at positions 700 and 715. Position 700 received a score of 0.292, higher than the score received by position 715 which was 0.208. These scores represent the probability that a protein from a data set of 172 proteins with known GPI-anchor addition sites has the same amino acid residues at positions ω, ω+1 and ω+2 as those predicted for PF08_0008.

There are four elements that are currently believed to constitute a GPI-anchor addition sequence (figure 1.3b) (Eisenhaber et al., 2001). Based on these rules, the ω site most strongly predicted by DGPI, position 700, is unlikely to be a true GPI addition site. Although there are two residues with small side chains (S at position ω and G at position ω+2) within the predicted cleavage site, the C-terminal hydrophobic tail begins 26 residues away from the ω site (figure 4.2b). The C-terminal hydrophobic tail must begin at around position ω+9 or ω+10. The presence of a ω site at residue 715 seems more likely, based on the sequence requirements. In this case there is an unstructured linker (ω-7 to ω-2), a cleavage site consisting of amino acids with small side chains (ω-1 to ω+2), a moderately polar spacer region (ω+3 to ω+10) and a hydrophobic region extending to the C-terminus (ω+11 to C-terminus).

The cleavage site predicted by Big-π also fits well with the sequence requirements (figure 4.2c). There is an unstructured linker (ω-6 to ω-2), a cleavage site consisting of amino acids with small side chains (ω-1 to ω+2), a moderately polar spacer region (ω+3 to ω+11), and a hydrophobic region extending to the C-terminus (ω+12 to C-terminus).

With both of these predictions there are a couple of slight variations from the sequence requirements detailed in figure 1.3b. The unstructured linker consists of 5 or 6 amino acids according to the Big-π or DGPI predictions respectively, whereas the described sequence requirements detail an unstructured linker with a length of approximately 11 amino acids. Also, the hydrophobic tail has been described to begin at positions ω+9 or ω+10, whereas according to the DGPI prediction it would start at position ω+11, and with the Big-π prediction, at ω+12.

The two programs used are not designed specifically for use with *P. falciparum* and are trained using animal, plant and fungal protein sequences. Although the sequence requirements for GPI addition are generally conserved between mammalian and protozoan proteins, they are not interchangeable. For instance it has been shown that the GPI addition sequence of the variant surface glycoprotein (VSG) of *Trypanosoma brucei* functions poorly in mammalian cells (Moran and Caras, 1994). Non-*Plasmodium*-trained algorithms are not ideal for use with the current query and may introduce false positive or negative results, based on these subtle differences. For example, the known GPI-anchored protein MSP1 was not predicted to be GPI-anchored by either the DGPI or Big-π algorithms.

With this in mind, another algorithm has been designed that has been trained using known or probable GPI-anchored *P. falciparum* proteins (Gilson et al., 2006), GPI-HMM. This algorithm was used to rank all proteins encoded in the *P. falciparum* genome according to their likelihood of being GPI-anchored. In the list of the top 30 most strongly predicted GPI-anchored proteins, PF08_0008 features at number 19. All of the proteins ranked above PF08_0008, apart from one, PF0975c, have been shown biochemically to be GPI-anchored, or are strongly predicted to be GPI-anchored based on homology with other known GPI-anchored proteins. Moreover, none of the proteins on this list has been shown not to be GPI-anchored.

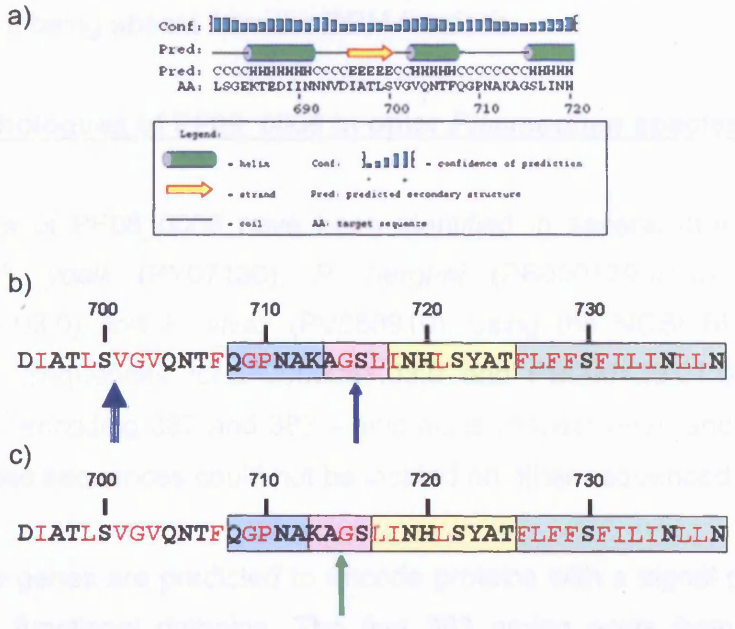
***Figure 4.2*: Prediction of GPI-anchorage by two algorithms, and GPI signal sequence requirements**

a) Prediction of the secondary structure of the extreme C-terminus of PF08_0008 by PSIPRED (McGuffin et al., 2000). It can be seen that the amino acids at positions 708 to 715 are predicted to form a random coil and therefore lack any secondary structure.

The extreme C-terminus of PF08_0008 was analysed in order to determine if the four sequence elements typical of a GPI-anchor addition sequence (described in figure 1.3b), are indeed present. The amino acid sequence of the extreme C-terminus is shown in b) and c), with hydrophobic residues coloured in red.

b) The thick blue arrow indicates the cleavage site (between residues ω and ω+1) confidently predicted by the DGPI algorithm, and the thin blue arrow indicates the second cleavage site predicted less confidently by this algorithm. The boxes enclose the four GPI-anchor signal sequence elements of the second cleavage site: unstructured linker (blue), cleavage site (pink), moderately polar spacer region (yellow), and hydrophobic tail (green).

c) As for b), except the green arrow indicates the cleavage site predicted by the Big-π algorithm, although this prediction was below the confidence threshold.



4.5 No evidence for GPI-anchorage of PF08_0008 in 'GPI proteome' studies

In the same study that resulted in the development of GPI-HMM, parasites were metabolically labelled with radioactive glucosamine and a proteomic analysis was carried out, in order to identify GPI-anchored proteins by 2-dimensional (2-D) gel electrophoresis followed by mass spectrometry. PF08_0008 was not positively identified. However, there were a number of spots on the 2-D gels that could not be identified by mass spectrometry, and it remains a possibility that PF08_0008 could be represented in one of those. In addition, it was estimated that the 11 biochemically confirmed GPI-anchored proteins represent around 94% of the GPI-anchored proteome. Other novel GPI-anchored proteins are therefore likely to be poorly abundant relative to these 11, and this could hinder their identification by conventional means.

In a similar study (Sanders et al., 2005), detergent-resistant membranes (DRMs) of mature schizonts were sucrose-gradient purified, since previous work had suggested GPI-anchored *P. falciparum* proteins are enriched in such membrane fractions (Black et al., 2003). The proteome of these DRMs was determined and found to contain the GPI-anchored proteins MSP1, MSP2, MSP4 and RAMA, as well as a number of novel putative GPI-anchored proteins. PF08_0008 was not detected in the DRM proteome. Again, this could be a consequence of PF08_0008 being poorly abundant relative to other GPI-anchored molecules, rather than it being absent from the DRM fractions.

4.6 Orthologues of PF08_0008 in other *Plasmodium* species

Orthologues of PF08_0008 have been identified in several other *Plasmodium* species: *P. yoelii* (PY07130), *P. berghei* (PB000129.01.0), *P. chabaudi* (PC000051.03.0) and *P. vivax* (PV088910), using the NCBI BLAST tool and PlasmoDB. Sequences for PC000051.03.0 and PB000129.01.0 are currently incomplete (encoding 387 and 383 amino acids respectively), and the remaining parts of these sequences could not be located on other sequenced contigs.

All of these genes are predicted to encode proteins with a signal peptide and no identifiable functional domains. The first 383 amino acids from the *P. yoelii*, *P. chabaudi* and *P. berghei* sequences were initially used in a ClustalW multiple

alignment (Thompson et al., 1994) to calculate similarity between these sequences. The results are shown in figure 4.3a i and it can be seen that there is likely to be a very high level of conservation between these species, based on the results for the available sequence.

The complete sequence of *P. yoelii* was therefore used as a representative rodent malaria parasite for a ClustalW (Thompson et al., 1994) multiple alignment with the entire *P. falciparum* and *P. vivax* sequences (figure 4.3b). The level of identity between the *P. falciparum* protein sequence and the sequences of its orthologues was found to be around 40% (figure 4.3a ii).

This existence of such orthologues is suggestive of a conserved role for this protein in *Plasmodium* species. Interestingly, 8 out of 9 of the cysteine residues in the *P. falciparum* protein sequence, which may be important structurally for the formation of disulphide bonds, are conserved in both PY07130 and PV088910.

***Figure 4.3*: Alignment of PF08_0008 with orthologues from other *Plasmodium* species**

a) i) ClustalW multiple alignments were used to calculate sequence identity between the three PF08_0008 orthologues in rodent malaria species.
ii) ClustalW multiple alignments were used to calculate sequence identity between PF08_0008, the *P. vivax* orthologue PV088910 and the *P. yoelii* orthologue PY07130.

b) A ClustalW multiple alignment of the PF08_0008, PV088910, and PY07130 protein sequences. PF08_0008 and PV088910 each contain nine cysteine residues and PY07130 contains 10. Eight of these are positionally conserved between all three species and are indicated with a yellow star. The remaining non-conserved cysteine residues are indicated with red crosses. Identical residues are coloured in pink, similar residues are coloured in blue.

c) Two ClustalW multiple alignments were carried out on regions of PF08_0008 and its orthologues that do not contain any low complexity regions. The alignments were used to calculate the sequence identity between the three sequences.

i) Residues 65-338 of the PF08_0008 sequence were aligned with the corresponding sequence from the *P. vivax* and *P. yoelii* orthologues, as aligned in b)
ii) Residues 500-653 of the PF08_0008 sequence were aligned with the corresponding sequence from the *P. vivax* and *P. yoelii* orthologues, as aligned in b).

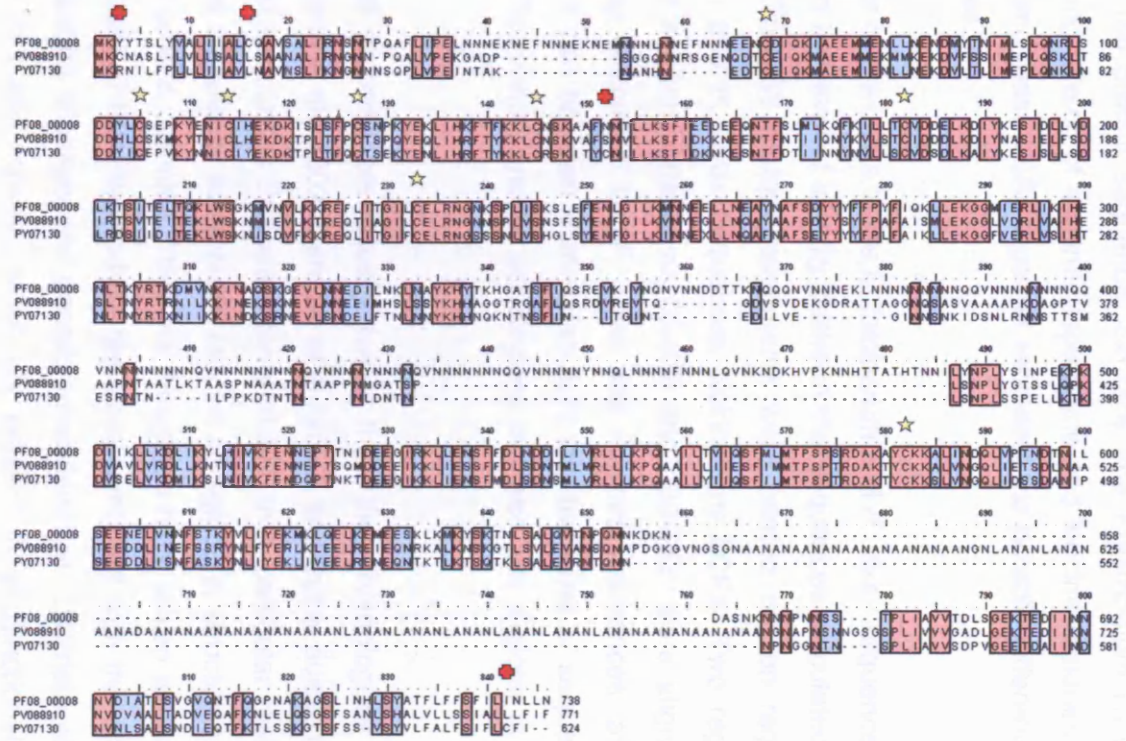
a) i)

	P. yoelii	*P. chabaudi*	*P. berghei*
P. yoelii	X	0.91	0.94
P. chabaudi	0.91	X	0.88
P. berghei	0.94	0.88	X

ii)

	P. falciparum	*P. vivax*	*P. yoelii*
P. falciparum	X	0.38	0.45
P. vivax	0.38	X	0.47
P. yoelii	0.45	0.47	X

b)



c) i)

	P. falciparum	*P. vivax*	*P. yoelii*
P. falciparum	X	0.59	0.59
P. vivax	0.59	X	0.67
P. yoelii	0.59	0.67	X

ii)

	P. falciparum	*P. vivax*	*P. yoelii*
P. falciparum	X	0.55	0.6
P. vivax	0.55	X	0.69
P. yoelii	0.6	0.69	X

4.7 PF08_0008 contains two asparagine-rich regions of low complexity

The alignment of PF08_0008 with PV088910 and PY07130 highlights an interesting feature of PF08_0008 - it contains two low complexity regions (LCRs), which are predominantly made up of asparagine residues. The first region encompasses amino acid residues 38-64 at the N-terminal end of the sequence. The second LCR is much larger and covers residues 356-485 towards the middle of the predicted protein sequence. Similarly, the *Plasmodium vivax* sequence contains a long stretch of alanine, asparagine and leucine residues, towards its C-terminus, and these LCRs together represent the notable differences between the sequences.

The presence of the LCRs in the *P. falciparum* and *P. vivax* sequences is likely to have affected the level of identity between the sequences calculated previously. Two more ClustalW multiple alignments were carried out on regions of the *P. falciparum* and *P. vivax* sequences lacking any LCRs. Two regions of the *P. falciparum* sequence (residues 65-338 and 500-653) were aligned with the corresponding regions of the *P. vivax* and *P. yoelii* sequences, as aligned in figure 4.3b. It can be seen from figure 4.3c that the level of sequence identity between *P. falciparum* and its orthologues in these two regions is around 55-60%.

P. falciparum proteins are usually larger than their homologues from other species (Florens et al., 2002, Gardner et al., 2002) and this is due to the insertion of amino acid sequences that are often enriched in a particular residue. These regions of low complexity are thought to form non-globular structures that extend from ordered protein domains. They are thought to have arisen as a result of the high A/T content of the *Plasmodium* genome, combined with the effects of this bias on replication slippage and genetic recombination (DePristo et al., 2006). The function of these repeats, which are present in all functional classes of *Plamodium* proteins, is unknown, although their presence in the *var* gene family is thought to facilitate intra-exonic recombination and hence antigenic diversity (DePristo et al., 2006).

4.8 PF08_0008 contains no predicted domains

A BLAST search using the PF08_0008 sequence does not yield any hits from characterised proteins of any known function, therefore these results cannot be used to imply any functional role for PF08_0008 itself. Other databases can be used to identify functionally important motifs within a query sequence that may be present in other proteins, even when the protein is distantly related to any other proteins of known function.

The Simple Modular Architecture Research Tool (SMART) database is designed to identify evolutionarily conserved protein domains within query sequences (Schultz et al., 1998). SMART contains information on over 500 domain families found in proteins with diverse cellular functions. When the PF08_0008 amino acid sequence is used to search the SMART database, no protein domain features are identified, except for the predicted signal peptide, asparagine-rich regions and C-terminal transmembrane domain, all of which have been identified through other methods.

ProSite is another searchable database that can be used to assign query sequences to protein families or to identify protein domains within the sequence. As was the case with SMART, no functional domains were identified in PF08_0008, nor could it be assigned to any protein family.

Notably, there does not appear to be a homologue of PF08_0008 in the related apicomplexan parasite *Toxoplasma gondii*, suggesting that it may be performing a function specific to *Plasmodium*.

4.9 Yeast 2-hybrid interactions

In 2005 a study was published in which a high-throughput version of the yeast two-hybrid assay was used to identify *P. falciparum* protein-protein interactions (LaCount et al., 2005). Two potential interactions involving PF08_0008 were identified. The interacting proteins were MSP7 and the hypothetical protein PF14_0495. In both cases, PF08_0008 was the prey.

The dataset of protein-protein interactions generated in this study was refined to remove false positives, for example interactions involving 'promiscuous' protein

fragments with many partners. However, false positives can arise for many different reasons, for example, plasmid mutations or non-specific interactions resulting from protein over-expression. It is therefore extremely likely that false positives constitute a significant proportion of the remaining dataset. Although PF14_0495 and MSP7 have similar expression profiles to PF08_0008 and are thus likely to be present at the same time within the parasite, suggestions of any interaction would need to be confirmed through experimental means such as co-immunoprecipitation.

4.10 Conclusions

4.10.1 Membrane anchorage of PF08_0008

The information generated from the various algorithms discussed above acts merely as a guide for future work - the nature of any membrane anchorage needs to be determined experimentally before any conclusions can be drawn.

However, the data discussed above suggests that PF08_0008 is likely to be GPI-anchored:

- Three of the four transmembrane domain prediction algorithms predicted a transmembrane domain at the extreme C-terminus.
- This was the sole prediction for TMHMM2. For the two algorithms that also predicted a transmembrane domain in the region of residues 555-575, the prediction at the C-terminus was stronger.

Based on the results for the four transmembrane prediction algorithms, it appears that the prediction at the extreme C-terminus is the strongest. However, the C-terminal hydrophobic region is more likely to be a GPI signal sequence than a true transmembrane domain, and this region does contain the typical GPI signal sequence elements. Importantly, the GPI-HMM algorithm, trained on *P. falciparum* GPI signal sequences, predicted PF08_0008 was likely to be GPI-anchored, although no evidence has been found in any proteomic or DRM studies.

4.10.2 Orthologues of *PF08_0008*

There are orthologues of *PF08_0008* in *P. vivax*, as well as in the rodent malaria parasites *P. yoelli*, *P. berghei*, and *P. chabaudi*. All of the cysteine residues of the *P. falciparum* sequence, excluding one in the predicted signal peptide, are conserved in the other *Plasmodium* species examined, suggesting the structure of PF08_0008 may be conserved. LCRs are found in the *P. falciparum* and *P. vivax* sequences, and it could be postulated that in both species these repeats are situated between structured or functional domains. Sequence identity between the different orthologues was found to be in the region of 55-60% in regions lacking LCRs. However, BLAST and conserved protein domain searches have failed to provide any information as to any possible function of this gene.

5. *Raising antibodies against N- and C-terminal regions of* PF08_0008

5.1 Introduction

In order to study PF08_0008 further, polyclonal antibodies were raised in mice and rabbits against different regions of the protein, with a view to allowing the characterisation of PF08_0008 through antibody-based techniques such as Western blot, immunofluorescence and immunoprecipitation.

Two regions of the PF08_0008 protein were chosen for recombinant expression, one from the N-terminus and one from the C-terminus. In the event that PF08_0008 is proteolytically processed, antibodies raised against separate regions of the protein would allow this to be studied more easily. The relative size and location of these regions within the PF08_0008 sequence, in relation to the predicted signal peptide and GPI-anchoring sequences are shown in figure 5.1.

5.2 Ligation-independent cloning (LIC)

Primers were designed to amplify two regions of the *PF08_0008* gene, corresponding to N- and C-terminal regions of the translated protein, incorporating ligation-independent cloning (LIC)-compatible overhangs into the PCR product. The expected sizes of the PCR products were 331 bp and 445 bp for the N- and C-terminal inserts respectively, including overhangs. PCR products were analysed by agarose gel electrophoresis (figure 5.1b).

The 331 bp N-terminal product was annealed to the pET 30 LIC vector, and the 445 bp C-terminal product to the pET 32 LIC vector. The pET 30 vector allows expression of recombinant protein fused with N-terminal His-tag and S-tag sequences for identification and purification purposes. The pET 32 vector also allows for expression of recombinant protein fused N-terminally with these two tags, in addition to a 109 amino acid thioredoxin protein tag to enhance solubility. *E. coli* One Shot TOP10 cells were transformed with these vectors, and plasmid DNA was obtained from a number of clones. Restriction digests confirmed the presence of the correctly-sized insert in the plasmid DNA (figure 5.1c).

***Figure 5.1*: Cloning two regions of *PF08_0008* into the expression vectors pET 30 and pET 32, by ligation-independent cloning (LIC)**

a) Two regions of PF08_0008 were chosen to be expressed as recombinant proteins, in order to generate antibodies. These regions are shown in green. The putative GPI-anchor signal sequence is shown in pink, and the N-terminal signal sequence in blue. The amino acid residues corresponding to the start and end of the N- and C-terminal regions chosen for recombinant expression are numbered, along with the extreme N- and C-termini of PF08_0008.

b) Specific primers were used to amplify two regions of *PF08_0008*, encoding the N- and C-terminal regions of the protein to be recombinantly expressed. The PCR products were analysed by agarose gel electrophoresis. Molecular sizes are given in bp.

i) The 331 bp fragment, encoding the N-terminal region

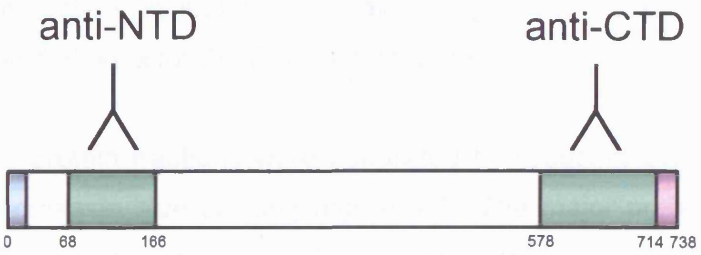
ii) The 445 bp fragment, encoding the C-terminal region

c) PCR products from b) were annealed to pET 30 (331 bp fragment) and pET 32 (445 bp fragment) vectors. *E. coli* One Shot TOP10 cells were transformed separately with the two plasmids, and DNA was obtained from resultant clones. Plasmid DNA was digested with *Bam*HI and *Bgl*II. These enzymes cut either side of the LIC site and should release the cloned fragment. Digested DNA was analysed by agarose gel electrophoresis. Molecular sizes are given in bp.

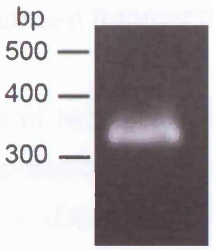
i) N-terminal region/pET 30 vector: Digested plasmid DNA from three clones numbered 1-3 is shown. Only clone 2 possesses the vector containing the correctly-sized insert (arrowed).

ii) C-terminal region/pET 32 vector: Digested plasmid DNA from three clones numbered 1-3 is shown. All three clones possess the vector containing the correctly-sized insert (arrowed).

a)


anti-NTD
anti-CTD
0
68
166
578
714 738

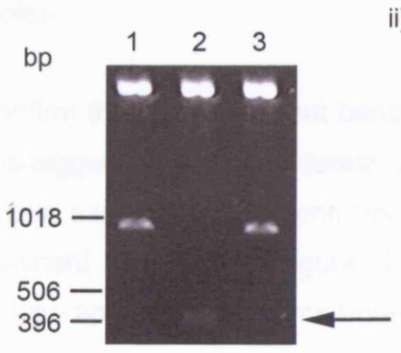
b) i)


bp
500
400
300

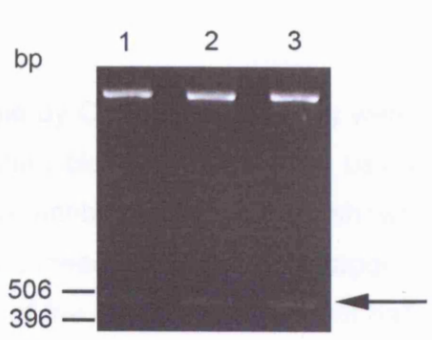
ii)


bp
506
396

c) i)


bp
1
2
3
1018
506
396

ii)


bp
1
2
3
506
396

5.3 Testing the solubility of recombinant protein

Plasmid DNA containing the correctly sized insert was gel-purified and used to transform *E. coli* BL21 (DE3) cells. Initially, small scale 1 ml cultures were grown to determine the solubility of the recombinant protein. The predicted molecular mass of recombinant protein, including the relevant tags, was 16 kDa for the N-terminal fragment and 33 kDa for the C-terminal fragment.

Soluble and insoluble protein fractions were separated by reducing SDS-PAGE and visualised by Coomassie blue staining (figure 5.2). The major protein band present in figure 5.2a i runs at approximately 40 kDa, and is present in the insoluble protein fraction of both the induced and un-induced samples. The major band in figure 5.2a ii runs at just under 18 kDa and again is present in the insoluble fraction of both induced and un-induced samples. These 40 kDa and 18 kDa bands are likely to correspond to the C-terminal and N-terminal recombinant protein fragments respectively.

The presence of recombinant protein in the un-induced pellet fraction suggests that 'leaky expression' of the recombinant proteins is occurring. In the presence of IPTG, BL21 (DE3) cells express the *lacUV5* promoter-controlled T7 RNA polymerase. The sequence encoding the protein of interest is cloned into the pET 30/32 vectors so that transcription is driven by the T7 RNA polymerase, and thus should only occur in the presence of IPTG. Transcription of the T7 polymerase gene in the absence of IPTG, and therefore transcription of the cloned gene, is likely to account for the presence of target protein in the un-induced samples.

In order to confirm that the prominent bands visible by Coomassie staining were in fact the His-tagged proteins of interest, a Western blot was carried out using the same protein samples and an anti-His primary antibody. Figure 5.2b shows that the prominent bands from figure 5.2a do indeed represent His-tagged proteins. It is also apparent that some breakdown of the C-terminal fragment has occurred.

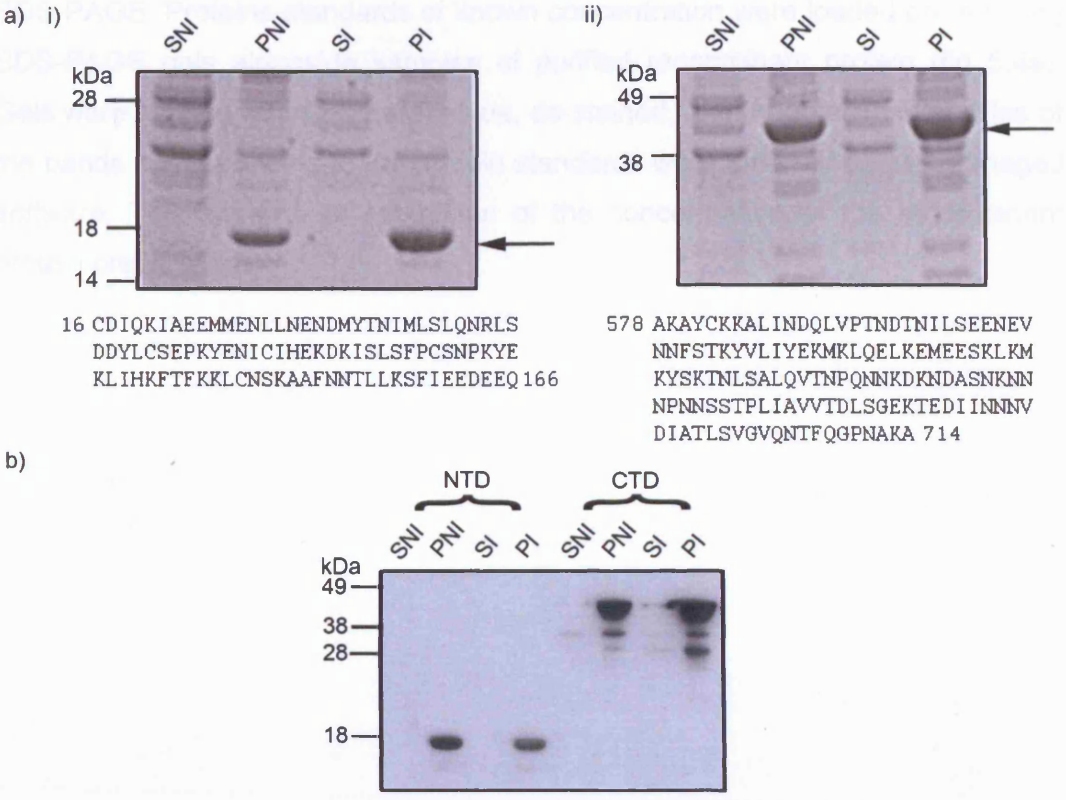
***Figure 5.2*: Expression of recombinant protein**

a) Coomassie-stained SDS-PAGE gels showing analysis of soluble and insoluble protein fractions from BL21 (DE3) cells, before and after induction with IPTG. Molecular mass is shown in kDa. SNI = soluble fraction, before induction. PNI = insoluble fraction, before induction. SI = soluble fraction, following IPTG induction. PI = insoluble fraction following IPTG induction. Bands representing the target recombinant protein of interest are arrowed. Amino acid sequences are shown.

i) Analysis of protein fractions from cells transformed with the pET 30 vector encoding the N-terminal fragment of PF08_0008 (predicted molecular mass 16 kDa)

ii) Analysis of protein fractions from cells transformed with the pET 32 vector encoding the C-terminal fragment of PF08_0008 (predicted molecular mass 33 kDa).

b) The same protein fractions as analysed in a) were used in Western blot analysis, using an anti-His primary antibody. The His-tagged recombinant proteins are detected in the insoluble protein fraction of both induced and un-induced cultures. It can be seen that there has been some C-terminal breakdown of the recombinant fragment expressed in pET 32.



5.4 Purification of recombinant protein

In order to obtain sufficient quantities of N- and C-terminal protein fragments for immunisation of animals, large scale cultures of BL21 (DE3) cells were grown. Recombinant protein was solubilised in a urea-containing buffer and passed down a pre-equilibrated nickel-nitrilotriacetic acid (Ni-NTA) resin column. The six consecutive histidine residues of the recombinant protein are able to specifically bind the chromatography matrix and the protein is immobilised. Washes with buffers containing low concentrations of imidazole, which can also bind to the matrix, were then carried out to disrupt the binding of dispersed histidine residues in non-tagged proteins. Recombinant protein was then eluted with 250 mM imidazole as the His-tagged protein is competed off the matrix. Samples of protein taken at each stage of the purification process were separated by SDS-PAGE and visualised by Coomassie blue staining (figure 5.3). It can be seen that both N- and C-terminal recombinant proteins were readily purified using this method.

Following extensive dialysis against PBS, the protein became insoluble, meaning that its concentration could not be established through conventional spectrophotometric means. Instead, the concentration was determined using SDS-PAGE. Proteins standards of known concentration were loaded on reducing SDS-PAGE gels alongside samples of purified recombinant protein (fig 5.4a). Gels were stained with Coomassie blue, de-stained, and the relative intensities of the bands corresponding to the protein standards were determined using ImageJ software. This allowed an estimation of the concentration of the recombinant protein preparations.

***Figure 5.3*: Purification of recombinant protein**

The insoluble protein fractions of BL21 (DE3) cultures expressing the recombinant N- and C-terminal domains of PF08_0008 were solubilised in a urea-containing buffer and passed down a pre-equilibrated nickel-nitrilotriacetic acid (Ni-NTA) resin column. The flow-through of unbound protein was collected in each case, and three wash steps were carried out with buffers containing increasing concentrations of imidazole. Recombinant proteins were eluted with buffer containing 250 mM imidazole. Samples of protein that passed through the column at each step were analysed by SDS-PAGE and Coomassie blue staining. Molecular mass is indicated in kDa.

FT = initial flow-through of unbound protein. 10 mM = wash step with buffer containing 10 mM imidazole. 20 mM = wash step with buffer containing 20 mM imidazole. 50 mM = wash step with buffer containing 50 mM imidazole. 250 mM = elution of recombinant protein with buffer containing 250 mM imidazole.

a) Purification of C-terminal domain, expressed using pET 32 vector.

b) Purification of N-terminal domain, expressed using pET 30 vector.

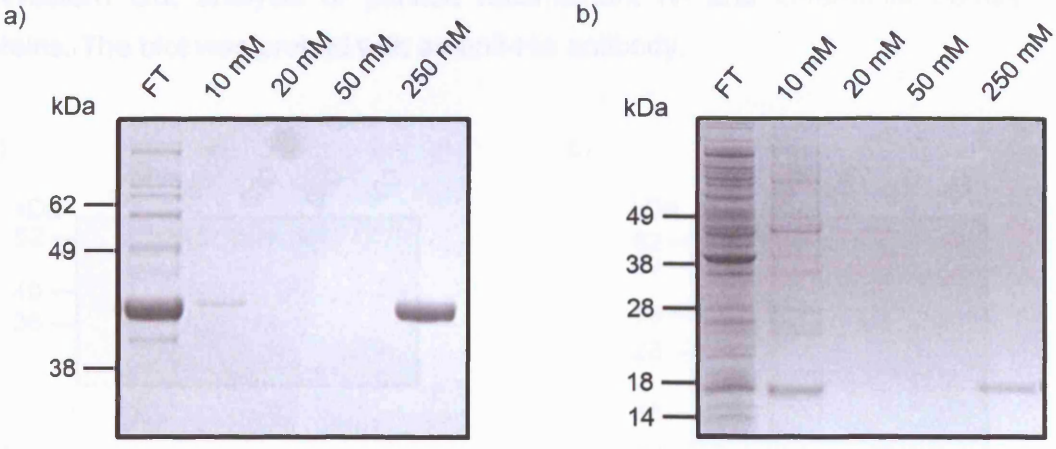


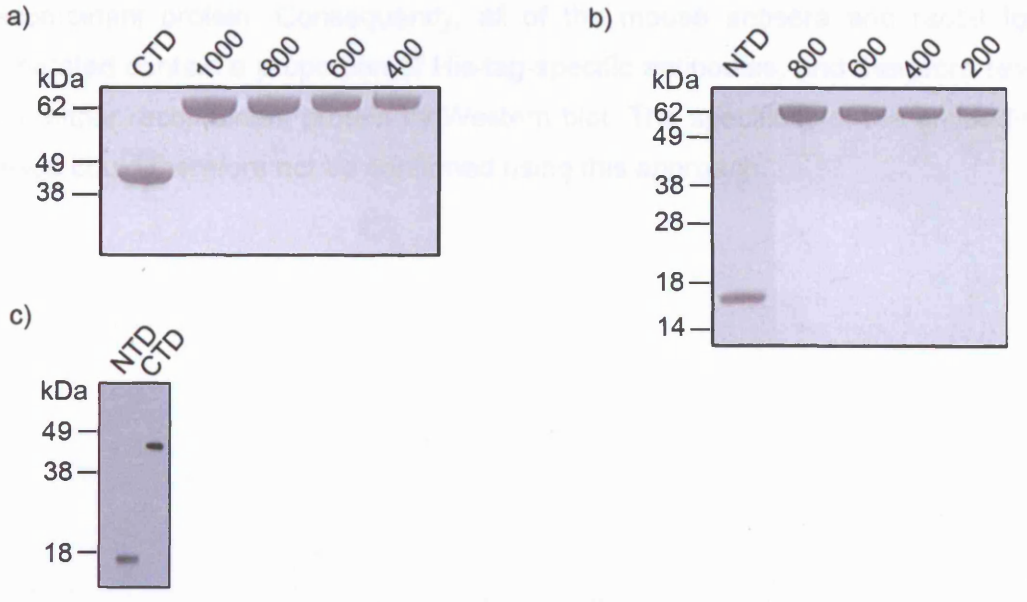
***Figure 5.4*: Quantification of purified recombinant protein**

The particulate nature of both recombinant proteins in PBS required protein concentration to be estimated by SDS-PAGE followed by Coomassie staining. Samples of recombinant protein were loaded alongside a range of protein standards of known protein concentration. Equal volumes of recombinant protein and protein standard were loaded.

a) Equal volumes of recombinant CTD protein were loaded alongside protein standards with known concentrations of 1000 µg ml^{-1}, 800 µg ml^{-1}, 600 µg ml^{-1}, and 400 µg ml^{-1}. Densitometry was performed on the resultant bands using the ImageJ program. A standard curve of band density against protein concentration was drawn using the data obtained for the protein standards (not shown). This allowed the calculation of the concentration of the protein sample, which in this instance was 609 µg ml^{-1}.

b) As a), but with equal volumes of recombinant NTD protein loaded alongside protein standards with known concentrations of 800 µg ml^{-1}, 600 µg ml^{-1}, 400 µg ml^{-1} and 200 µg ml^{-1}. The concentration of recombinant protein in this instance was calculated as 286 µg ml^{-1}.

c) Western blot analysis of purified recombinant N- and C-terminal domain proteins. The blot was probed with an anti-His antibody.



5.5 Production of polyclonal antibodies in mice and rabbits

As can be seen from figure 5.4a, the preparations of recombinant protein were very pure. Mice and rabbits were therefore immunised with the recombinant proteins without the need for further purification.

Sera from each of the 5 mice immunised per antigen were kept separate. In all of the experiments in which mouse anti-CTD antibodies were required, unprocessed mouse serum was used.

Immunoglobulin G (IgG) was purified from the serum of one mouse immunised with recombinant N-terminal protein, using a 1 ml HiTrap Protein G column. Figure 5.5a shows that the IgG eluted in fractions 3-7. These fractions were pooled, concentrated and buffer exchanged against PBS. This IgG preparation was used in all experiments where mouse anti-NTD antibodies were required.

Rabbit antibodies were received from Harlan as IgG and so no further purification was necessary. Rabbit pre-bleed sera were not purified by Harlan and so a HiTrap Protein G purification was carried out on one rabbit pre-bleed serum sample (see figure 5.5b). Fractions 5-10 were pooled and buffer exchanged against PBS.

Affinity purification of antisera would have been desirable, but this was not possible since most common antigen immobilisation techniques require soluble recombinant protein. Consequently, all of the mouse antisera and rabbit IgG generated contain a proportion of His-tag-specific antibodies, and therefore react with either recombinant protein by Western blot. The specificity of the antibodies raised could therefore not be confirmed using this approach.

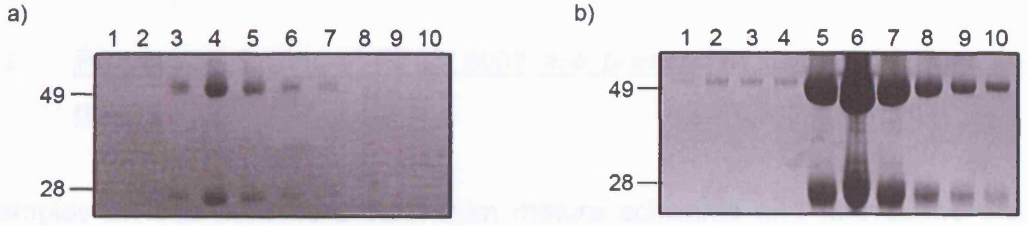
***Figure 5.5*: Purification of IgG from mouse antiserum and rabbit pre-immune serum**

1ml Protein G HiTrap columns were used to purify IgG from mouse antiserum raised against the N-terminal domain recombinant protein, and from rabbit pre-immune serum obtained from Harlan. IgG was eluted with 0.1 M glycine *p*H 2.7 into 10 0.5 ml fractions. Samples of all 10 fractions were analysed by SDS-PAGE followed by Coomassie staining.

SDS-PAGE analysis of the 10 elution fractions obtained following protein G purification of:

a) mouse antiserum, raised against the N-terminal domain recombinant protein

b) rabbit pre-immune serum.



6. PF08_0008 *is proteolytically processed within P. falciparum schizonts and shed into culture supernatant*

6.1 Introduction

A common feature of *P. falciparum* and *T. gondii* proteins involved in invasion is their proteolytic processing. For example, the TgMIC2/TgM2AP complex in *T. gondii* and the *Plasmodium* proteins AMA1, MSP1, and MTRAP are all proteolytically processed. This can occur during transport to their sub-cellular location, following their secretion onto the parasite surface, or both, and such processing events are believed to be critical in the functioning of these proteins (reviewed in (Dowse and Soldati, 2004)).

It was decided to investigate whether PF08_0008 is subject to any proteolytic cleavages.

6.2 Processed forms of PF08_0008 are present in both schizonts and merozoites

Samples of total parasite protein from mature schizonts and merozoites were analysed by Western blot, using the mouse anti-CTD antibody, and the results can be seen in figure 6.1a. This antibody detects 3 different protein species, with apparent molecular masses of 72 kDa, 49 kDa and 42 kDa. The same band pattern is observed whether schizont or merozoite material is analysed. Both the rabbit and mouse anti-NTD antibodies have proven to be of limited use in Western blot analyses when used on total parasite protein material. The reactivity of the antibodies with PF08_0008 relative to background parasite proteins is low, and this is presumably due to the presence of low levels of PF08_0008-specific antibodies within the serum.

In a separate experiment schizont material was treated with a high *p*H carbonate buffer, in order to enable separation of soluble or peripheral membrane proteins from those proteins with a membrane anchor. A Western blot carried out on the different fractions obtained in this extraction using the mouse anti-CTD antibody showed that the 72 kDa and 49 kDa forms of PF08_0008 are membrane-anchored (figure 6.1b). It is also evident that in this particular preparation of

parasite material, PF08_0008 is present almost exclusively in the 49 kDa form, with very little 72 kDa form and no visible 42 kDa form.

The 72 kDa protein that is detected by the anti-CTD antibody in schizonts and merozoites is presumably full-length PF08_0008, since the predicted molecular mass of the protein, minus the signal peptide, is 83 kDa (http://expasy.org/cgi-bin/pi_tool). The predicted molecular mass of the protein minus the signal peptide, and also the putative GPI signal sequence, is 80 kDa. This discrepancy between the observed and predicted molecular mass is likely to be due to the fact that molecular mass calculations were performed using molecular mass standards of 62 kDa and 188 kDa, with no intermediate, and therefore may not be accurate for values in this range.

The 49 kDa and 42 kDa bands, denoted p49 and p42, presumably represent products of two processing events. When the size of the processing products, the location and length of the region against which the C-terminal antibodies were raised (figure 6.1c), and the fact that full-length PF08_0008 and p49 appear to be membrane anchored are taken into account, it seems probable that any proteolytic cleavages occurring are in the protein ectodomain, N-terminal of the region expressed to raise C-terminal antibodies.

***Figure 6.1*: Western blot analysis of PF08_0008**

a)

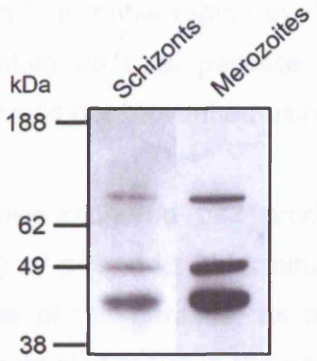
i) Analysis of the reactivity of anti-CTD antibodies with *P. falciparum* schizont and merozoite material separated under reducing conditions. Three polypeptides were detected in each sample.

ii) The distances travelled along the gel by the protein standards were measured and used to plot a standard curve of distance against $\log_{10}$(molecular mass). Data points are connected in black, and the trendline displayed in red. This enabled the calculation of the molecular mass of the three polypeptides, using their migration distances and the equation displayed. The R^2 value is also indicated on the graph. The table shows the data used to produce the standard curve. The values marked with an asterisk represent data for the three polypeptides.

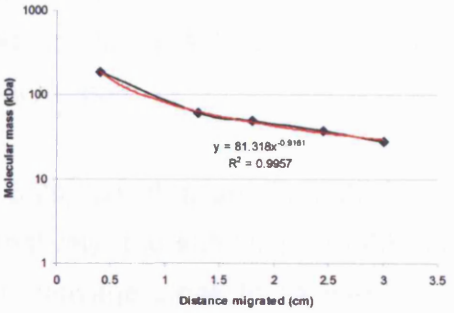
b) Analysis of the reactivity of anti-CTD, anti-PTRAMP and anti-MSP3 antibodies with high *p*H carbonate buffer-treated *P. falciparum* schizont material. Total parasite material (T) and fractions of parasite material soluble (CS) or insoluble (CP) in 100 mM Na_2CO_3 buffer (*p*H 11.0) were separated under reducing conditions. The 72 kDa and 49 kDa polypeptides of PF08_0008 are detected in total parasite and carbonate-insoluble material. The type I transmembrane protein PTRAMP is detected in the total and carbonate-insoluble material, and the soluble merozoite surface protein MSP3 in the total and carbonate-soluble material, both as expected.

c) Schematic representation of the size, and location within PF08_0008, of recombinant protein fragments expressed and used to generate antibodies against the N- and C-terminal regions. Location of the signal peptide (blue) and putative GPI-anchorage sequence (pink) are also shown.

a) i)

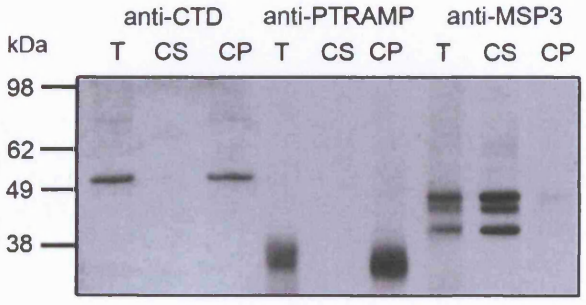


ii)



Distance migrated (cm)	Molecular mass (kDa)
0.4	188
1.3	62
1.8	49
2.45	38
3	28
*1.15	72
*1.75	49
*2.05	42

b)



c)

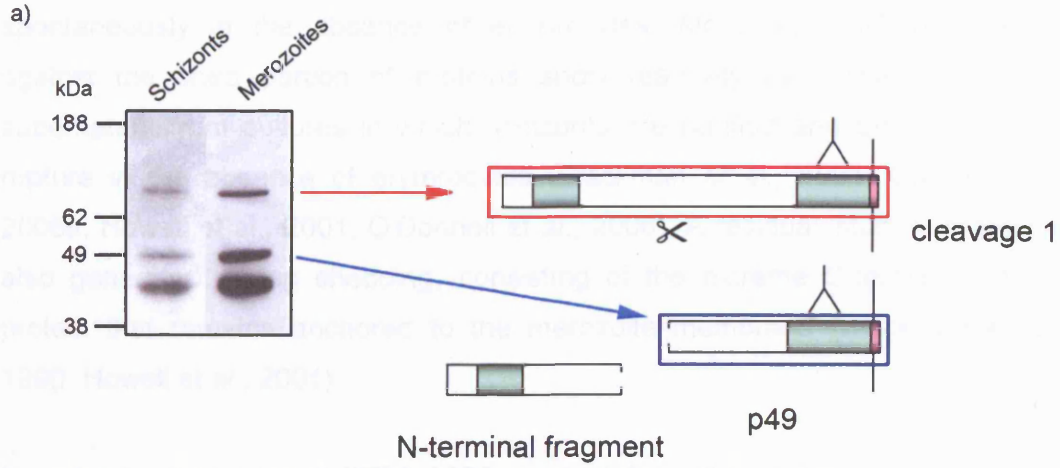
Schematic representations of cleavages which result in the generation of p49 and p42 are shown in figure 6.2. In order for the C-terminal p49 fragment to be generated (figure 6.2a), an N-terminal by-product of around 32 kDa would also be generated, assuming a molecular mass of 80 kDa. However, antibodies raised against the N-terminal region of PF08_0008 have not been used successfully in Western blots on total parasite material and so the existence of a 32 kDa fragment could not be verified using this approach alone.

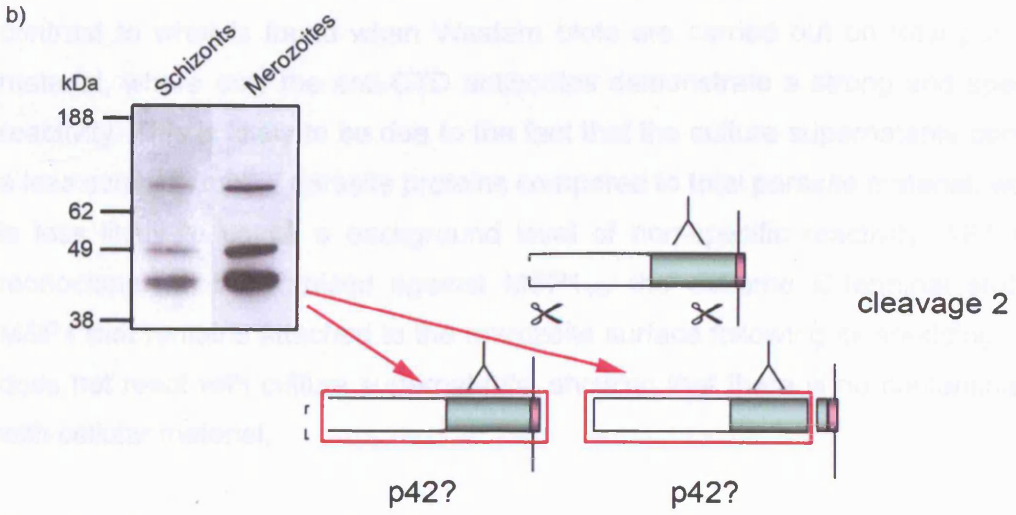
A membrane-anchored p42 product (figure 6.2b) could arise through further processing of p49 at its N-terminal end. Alternatively, the shedding of p49 from the surface of the parasite, as a result of a cleavage close to its membrane anchor, could also result in the generation of soluble p42. The absence of p42 in the parasite preparations used for the extraction of membrane proteins (figure 6.1b) has made it difficult to confidently predict which cleavage event occurs.

***Figure 6.2*: Schematic representations of proteolytic cleavages that could result in the generation of p49 and p42**

a) Full-length PF08_0008 is shown schematically on the right of the figure, membrane-anchored at its C-terminus, enclosed in a red box. The top band observed in the Western blot analysis (left) corresponds to this full length protein. This is cleaved to give the C-terminal p49 fragment, enclosed in blue, which corresponds to the middle band on the Western analysis. An N-terminal product is also generated, but is not detected by anti-CTD antibodies.

b) p42 could be generated by either of two cleavage events. p49 could be cleaved near to its membrane anchor, resulting in the 'shedding' of p42 from the merozoite surface. Alternatively, further processing of p49 at its N-terminus could result in the generation of a membrane-bound p42.


a)
kDa
Schizonts
Merozoites
188
62
49
38
cleavage 1
p49
N-terminal fragment


b)
kDa
Schizonts
Merozoites
188
62
49
38
cleavage 2
p42?
p42?

6.3 PF08_0008 is shed into culture supernatant

Some parasite molecules that are important in the invasion process such as AMA1, MSP1 and EBA175 are quantitatively shed from the surface of the merozoite, through two distinct mechanisms involving different classes of enzyme, as the parasite invades (reviewed in (O'Donnell and Blackman, 2005)). Such 'shedding' of molecules from the surface of merozoites also happens spontaneously in the absence of erythrocytes. Moreover, antibodies raised against the shed portion of proteins show reactivity by Western blot with supernatant from cultures in which schizonts are purified and then allowed to rupture in the absence of erythrocytes (Blackman et al., 1991, Green et al., 2006a, Howell et al., 2001, O'Donnell et al., 2006). A residual 'stub' of protein is also generated during shedding, consisting of the extreme C-terminus of the protein that remains anchored to the merozoite membrane (Blackman et al., 1990, Howell et al., 2001).

In order to determine if PF08_0008 is shed from the surface of merozoites, Western blotting was carried out under reducing conditions on culture supernatant. The results are shown in figure 6.3a. It can be seen that both the anti-CTD and anti-NTD antibodies are reactive with culture supernatant. This is in contrast to what is found when Western blots are carried out on total parasite material, where only the anti-CTD antibodies demonstrate a strong and specific reactivity. This is likely to be due to the fact that the culture supernatants contain a less complex mix of parasite proteins compared to total parasite material, which is less likely to cause a background level of non-specific reactivity. 1E1 is a monoclonal antibody raised against $MSP1_{19}$, the extreme C-terminal stub of MSP1 that remains attached to the merozoite surface following its shedding. 1E1 does not react with culture supernatants, showing that there is no contamination with cellular material.

The anti-CTD antibody recognises a doublet of proteins running at approximately 43 kDa. This is consistent with one of the models discussed above, in which the p49 fragment is cleaved near to its membrane anchor. However, the presence of a doublet suggests that such a cleavage takes place at two alternative sites. It is entirely possible that molecules that are shed from merozoites into culture supernatant, either spontaneously or upon productive invasion, are present in the purified schizont and merozoite preparations used for Western blot analysis.

Therefore, given their near-identical apparent size, it is likely that the 43 kDa doublet detected by the anti-CTD antibodies in culture supernatant is the same as the p42 fragment detected in total schizont and merozoite material, as shown in figure 6.1.

The anti-NTD antibody recognises a polypeptide with an apparent molecular mass of 37 kDa in culture supernatant. Since it is expected that an N-terminal fragment of approximately 32 kDa would be produced simultaneously with p49, it is possible that the 37 kDa fragment detected in culture supernatant is in fact this N-terminal product. This N-terminal fragment could be shed independently of the p42 shedding, or simultaneously, perhaps as part of a complex.

The detection of C-terminal fragments of PF08_0008 in culture supernatant also has implications regarding the membrane anchorage of this protein (figure 6.3b). The region of PF08_0008 between, but not including, the putative transmembrane and GPI-anchoring sequences was expressed recombinantly and used to generate the anti-CTD antibodies. It is clear from Western analysis detailed above that the anti-CTD antibodies detect fragments of PF08_0008 in culture supernatant. If the protein was anchored via a transmembrane domain, these antibodies would not detect any shed fragments, since their specificity lies C-terminal of the putative transmembrane domain. Similarly, antibodies raised against the C-terminal end of the AMA1 cytoplasmic region do not react with culture supernatant, whereas antibodies raised against the AMA1 ectodomain are able to detect the various shed fragments (Howell et al., 2005). However, anchorage of PF08_0008 via a GPI-anchor at the extreme C-terminus could allow shed fragments containing the antigenic region to be released into culture supernatant.

***Figure 6.3*: Western blot analysis of *P. falciparum* culture supernatant**

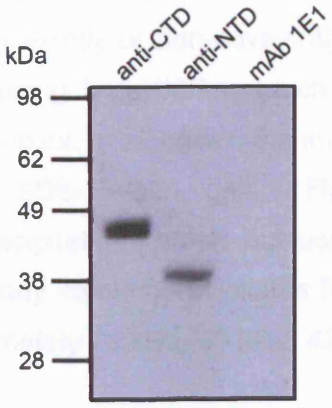
a)

i) Samples of supernatant were separated by SDS-PAGE under reducing conditions, transferred to nitrocellulose and probed with antibodies raised against the N- and C-terminal regions of PF08_0008, as well as the anti-$MSP1_{19}$ monoclonal antibody 1E1. Anti-CTD antibodies detect a doublet of polypeptides and anti-NTD antibodies detect a single polypeptide. No reactivity is exhibited by 1E1, a monoclonal antibody raised against the extreme C-terminal stub of MSP1 ($MSP1_{19}$) that remains attached to the merozoite surface following shedding of MSP1.

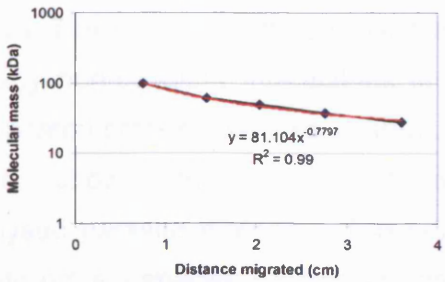
ii) The distances travelled along the gel by the protein standards were measured and used to plot a standard curve of distance against $\log_{10}$(molecular mass). Data points are connected in black, and the trendline displayed in red. This enabled the calculation of the molecular mass of the two polypeptides shown in i), using their migration distances and the equation displayed. The R^2 value is also indicated on the graph. The table shows the data used to produce the standard curve. The values marked with an asterisk represent data for the two polypeptides.

b) Schematic showing the location of the region of anti-CTD antibody specificity (green) on PF08_0008 in both the transmembrane and GPI-anchored models of protein anchorage. Hydrophobic sequences acting as potential transmembrane or GPI addition sequences are shown in pink. The parasite plasma membrane is indicated by a vertical line.

a) i)

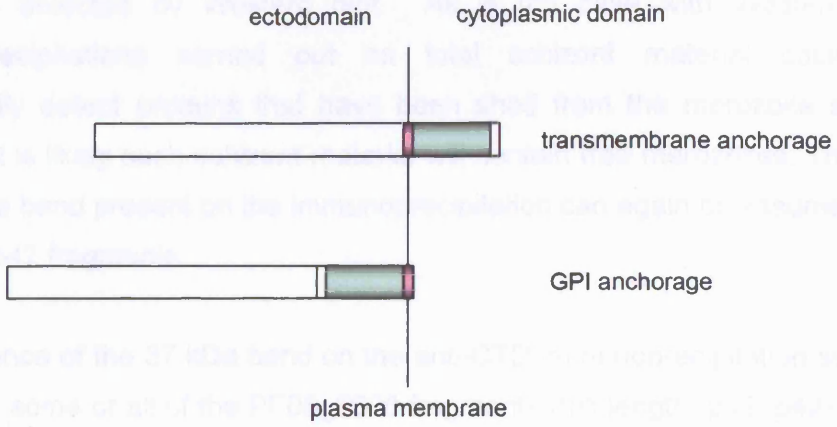


ii)



Distance migrated (cm)	Molecular mass (kDa)
0.75	98
1.45	62
2.03	49
2.75	38
3.6	28
*2.25	43
*2.7	37

b)



6.4 Anti-CTD antibodies immunoprecipitate multiple protein species from NP40- and SDS-solubilised ^{35}S-labelled parasites

Anti-CTD and anti-NTD antibodies were used in immunoprecipitation experiments carried out on extracts of ^{35}S-labelled mature schizonts (42 hours post-invasion (hpi)). Two different solubilisation methods were used. The first involved lysing parasite material in a buffer containing the non-ionic detergent NP40. This treatment allows the solubilisation and immunoprecipitation of protein complexes that are covalently or non-covalently associated. The second method utilised an SDS-containing lysis buffer, which disrupts any non-covalent interactions in a protein complex. In all cases, the immunoprecipitated proteins were separated on reducing SDS-PAGE gels. Figure 6.4a shows the result of an immunoprecipitation carried out using NP40-lysed parasite material. The anti-CTD antibody immunoprecipitates four separate protein species running at sizes of approximately 72 kDa, 49 kDa, 42 kDa and 37 kDa.

The immunoprecipitation band pattern obtained with the anti-CTD antibody is strikingly similar to the pattern obtained when this antibody is used in a Western blot (figure 6.1), with one notable exception - the presence of the additional 37 kDa band. Therefore it can be assumed that the 72 kDa and 49 kDa bands obtained by immunoprecipitation represent the full-length protein and p49 fragments detected by Western blot. As is the case with Western blots, immunoprecipitations carried out on total schizont material could also theoretically detect proteins that have been shed from the merozoite surface, because it is likely such schizont material will contain free merozoites. Therefore the 42 kDa band present on the immunoprecipitation can again be assumed to be the shed p42 fragments.

The presence of the 37 kDa band on the anti-CTD immunoprecipitation suggests that either some or all of the PF08_0008 fragments (full-length, p49, p42) exist in a complex with an unrelated 37 kDa protein. Alternatively, this 37 kDa protein could represent the N-terminal product that arises following full-length PF08_0008 processing to p49.

Figure 6.4b shows that when an immunoprecipitation is carried out on NP40-lysed parasite material using the anti-NTD antibodies, a very similar pattern of bands is produced on the autoradiograph as with the anti-CTD antibody. The only

inconsistency surrounds the presence or absence of the p42 product. The quality of the autoradiograph makes it difficult to determine whether a band is present or not and unfortunately the anti-NTD antibodies have proven to be of limited use in further immunoprecipitation experiments. Relative to the results obtained with the anti-CTD antibodies, very faint bands are produced on autoradiographs, making it hard to draw objective conclusions.

When an immunoprecipitation with the anti-CTD antibody is carried out following the disruption of non-covalently associated protein complexes with an SDS-containing lysis buffer, the same protein species are immunoprecipitated as when NP40-buffer is used (figure 6.5). Full-length protein, p49, p42 and the 37 kDa fragments are visible in both the SDS and NP40 lanes. The bands in the SDS lane appear less intense than in the NP40 lane, in particular the band corresponding to full-length PF08_0008, but they are clearly all still present. This result suggests that the 37 kDa protein species, whether it is a fragment of PF08_0008 or not, is linked to either the full length PF08_0008, its C-terminal processing fragments, or both, by one or more disulphide bonds.

***Figure 6.4*: Immunoprecipitation of PF08_0008 from NP40-lysed schizonts**

Immunoprecipitation of ^{35}S-labelled proteins from NP40-lysed mature schizont material with PF08_0008 mouse anti-CTD (a) and anti-NTD (b) antibodies. No proteins were immunoprecipitated with normal mouse serum (not shown).

Anti-CTD and anti-NTD antibodies both immunoprecipitate full-length PF08_0008, the p49 polypeptide and the 37 kDa fragment. The anti-CTD antibodies also immunoprecipitate p42, although the quality of the autoradiograph makes it difficult to determine if p42 is also immunoprecipitated by the anti-NTD antibodies.

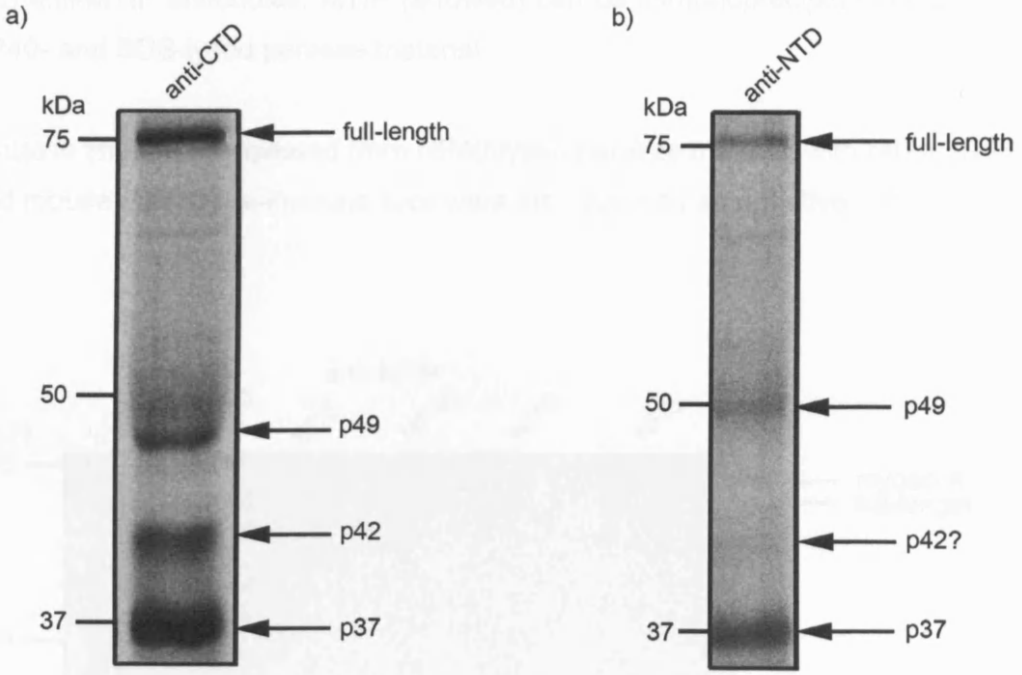
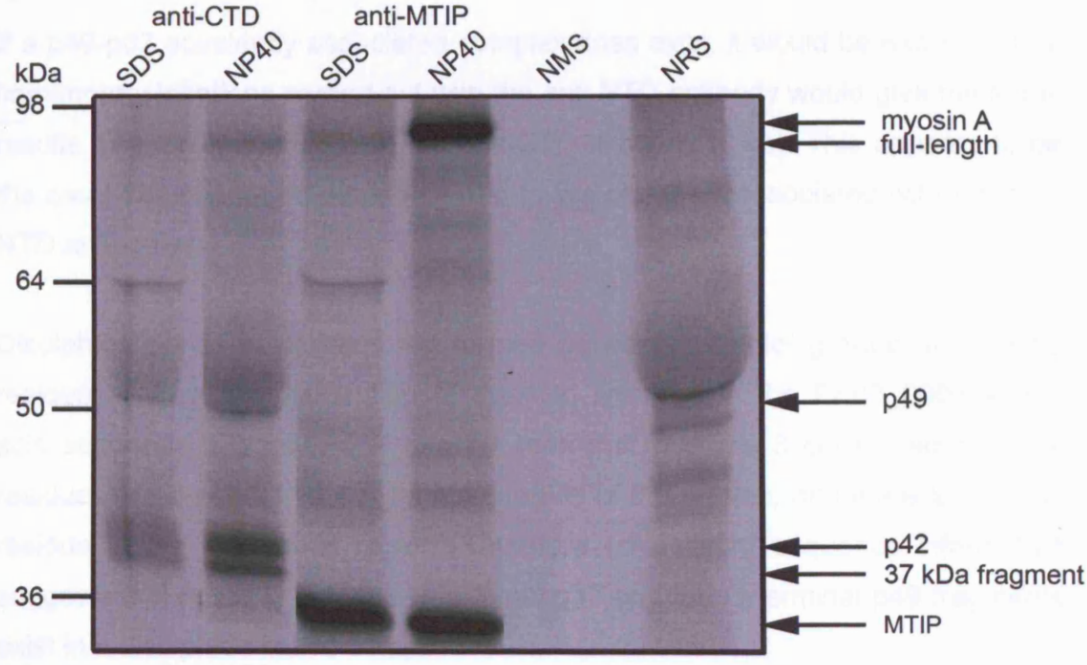


***Figure 6.5*: Immunoprecipitation of PF08_0008 from SDS-treated schizonts**

Immunoprecipitation with anti-CTD antibodies of ^{35}S-labelled proteins from mature schizonts, lysed with NP40 lysis buffer and an SDS-containing lysis buffer. All polypeptides precipitated from NP40-lysed parasite material can also be immunoprecipitated from SDS-lysed material (polypeptides arrowed and labelled full length, p49, p42 and 37 kDa fragment).

Myosin tail interacting protein (MTIP) is present in a complex with myosin A (Green et al., 2006b), although this complex can be disrupted with SDS. Anti-MTIP antibodies were used as a positive control for non-covalent complex disruption by SDS. Myosin A (arrowed), with an apparent molecular mass of ~92 kDa, can be immunoprecipitated from NP40- but not SDS-lysed parasite material with anti-MTIP antibodies. MTIP (arrowed) can be immunoprecipitated from both NP40- and SDS-lysed parasite material.

Proteins immunoprecipitated from NP40-lysed parasite material with rabbit (NRS) and mouse (NMS) pre-immune sera were also included as negative controls.



6.5 PF08_0008 processing products remain covalently associated

An immunoprecipitation was carried out on ^{35}S-labelled NP40-lysed parasite material using the anti-CTD antibody. The precipitated proteins were separated, in duplicate, by reducing SDS-PAGE. One set of samples was transferred to nitrocellulose and probed with the anti-NTD antibody by Western blot. The other sample was exposed to film as normal. This was in order to investigate the possibility that the 37 kDa protein species immunoprecipitated with the anti-CTD antibody is actually the N-terminal fragment that is generated following proteolytic cleavage of full-length protein.

Figure 6.6a shows that when used to probe proteins immunoprecipitated by the C-terminal antibody, the N-terminal antibody detects a band at approximately 40 kDa. This suggests that the 37 kDa band observed on autoradiographs following immunoprecipitation with the anti-CTD antibody is indeed the N-terminal portion of PF08_0008, hereby denoted p37. If no reactivity had been seen with the anti-NTD antibody on the blotted immunoprecipitation products, it would be suggestive of an interaction with an unrelated protein. If the result shown in figure 6.5 is taken into account, it can be deduced that the p49-p37 complex is held together by disulphide bonds.

If a p49-p37 covalently associated complex does exist, it would be expected that immunoprecipitations carried out with the anti-NTD antibody would give the same results as those obtained with the anti-CTD antibody (6.4b). This appears to be the case, but cannot be confirmed due to the problems associated with the anti-NTD antibodies.

Disulphide bonds in proteins are formed between the thiol groups of cysteine residues. When the distribution of cysteine residues in the PF08_0008 amino acid sequence is analysed, it can be seen that 7 of the 8 conserved cysteine residues are present in the N-terminal region of the protein, and there is a single residue in the C-terminal region. Therefore amino acid sequence information suggests it is possible that the N-terminal p37 and the C-terminal p49 fragments exist in a disulphide-linked complex.

Further support of a p49-p37 complex comes from performing Western blots on total schizont material under non-reducing conditions. A disulphide linkage

between the two fragments would be preserved under these conditions, and so it would be expected that the complex would have an apparent molecular mass similar to that of the full length PF08_0008, assuming no significant part of the sequence is lost during cleavage. Figure 6.6c shows that this is indeed the case. On a Western blot carried out under non-reducing conditions, the anti-CTD antibodies detect two protein species with very similar apparent sizes, in the region of 72 kDa. These are likely to represent full length PF08_0008, and the p49-p37 complex. The difference between the size of the two polypeptides could be due to the loss of part of the PF08_0008 sequence at the point of cleavage.

In order to confirm that the p42 and p37 fragments detected in culture supernatant are shed as a complex, Western blot analysis was carried out on culture supernatant under non-reducing conditions (figure 6.6d). If p42 and p37 are indeed shed as a complex, it would be expected that under non-reducing conditions, the anti-CTD and anti-NTD antibodies would detect protein species with a greater molecular mass than under reducing conditions, a mass that is slightly less than that of the full-length protein. It can be seen that this is indeed the case, with each antibody detecting at least two protein species. The multiple bands are likely to be due to the alternative shedding cleavages which are responsible for the generation of the doublet of p42 fragments that are seen under reducing conditions.

***Figure 6.6*: Western blot analysis of immunoprecipitated proteins**

a) Proteins immunoprecipitated from ^{35}S-labelled NP40-lysed schizonts with anti-CTD antibodies were separated by SDS-PAGE, transferred to nitrocellulose and probed with anti-NTD antibodies. Molecular masses are shown in kDa.

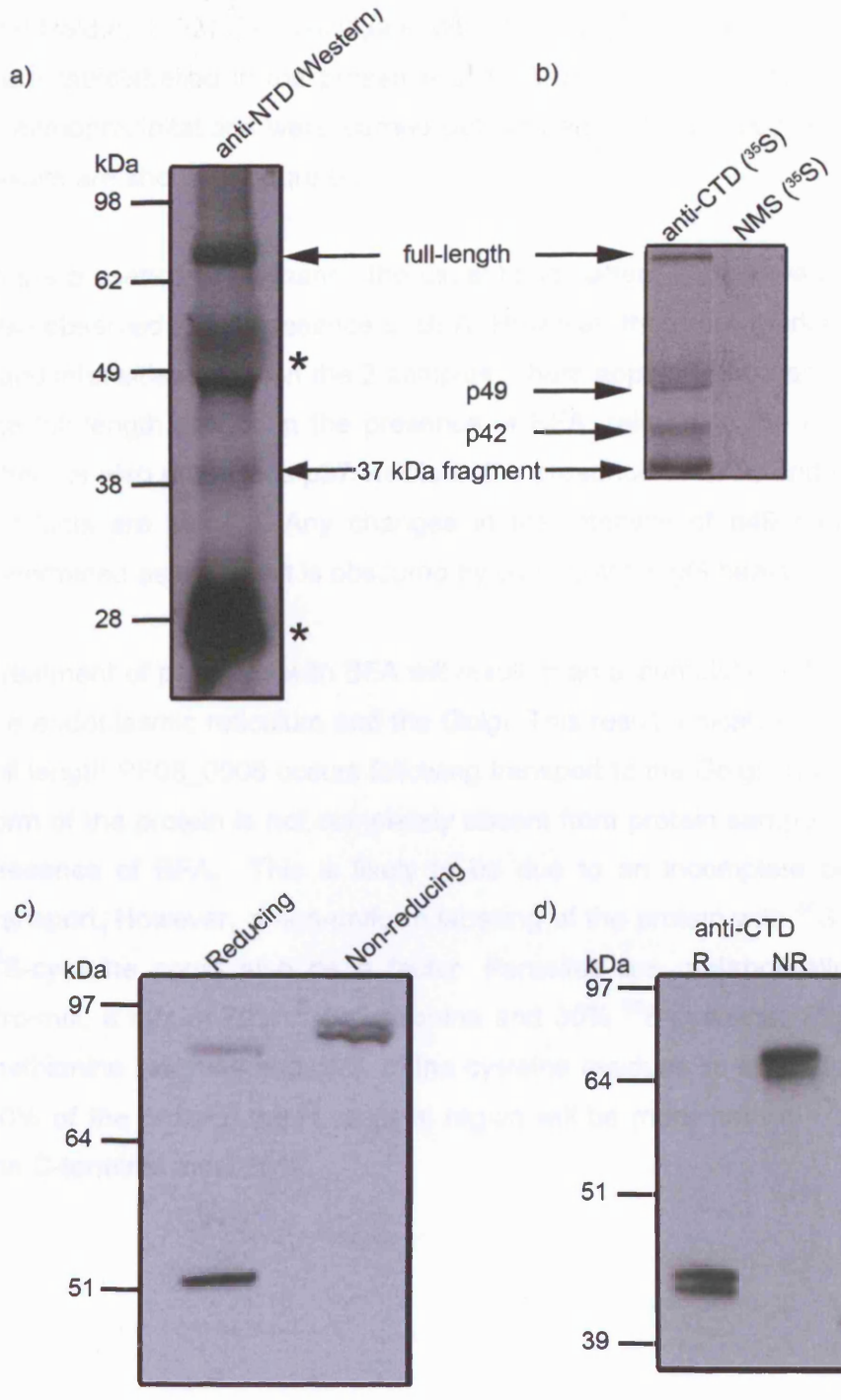
b) Autoradiograph showing proteins immunoprecipitated from the same preparation of schizonts used in a), using anti-CTD antibodies (anti-CTD (^{35}S)) or mouse pre-immune serum (NMS (^{35}S)). Molecular masses are shown in kDa.

Both a) and b) are on the same scale.

No proteins were immunoprecipitated by the mouse pre-immune serum, suggesting that all bands immunoprecipitated by the anti-CTD antibodies are specific PF08_0008 polypeptides or binding partners. Anti-NTD antibodies react by Western blot with both the full length PF08_0008 and the 37 kDa polypeptide that are immunoprecipitated by the anti-CTD antibodies (both polypeptides arrowed and labelled), suggesting that the 37 kDa fragment is indeed the N-terminal product of full length PF08_0008 processing. Mouse anti-NTD antibodies also react by Western blot with the heavy and light chains of the anti-CTD IgG antibodies present in the sample (marked with asterisks). The anti-NTD antibodies do not react with the 42 kDa polypeptide immunoprecipitated by the anti-CTD antibodies, and presumably not the 49 kDa polypeptide, although the latter is hard to confirm due to reactivity with immunoglobulin heavy chain.

c) Western blot of total schizont material separated by SDS-PAGE under reducing and non-reducing conditions. Blots were probed with mouse anti-CTD antibodies. Molecular masses are shown in kDa.

d) Western blot of culture supernatant separated by SDS-PAGE under reducing (labelled R) and non-reducing (labelled NR) conditions, and probed with mouse anti-CTD and anti-NTD antibodies. Molecular masses are shown in kDa.


a)
anti-NTD (Western)
kDa
98
62
49
38
28
full-length
*
*
37 kDa fragment
b)
anti-CTD (^{35}S)
NMS (^{35}S)
p49
p42
c)
Reducing
Non-reducing
kDa
97
64
51
d)
anti-CTD
anti-NTD
R
NR
R
NR
kDa
97
64
51
39

6.6 Processing of PF08_0008 is sensitive to brefeldin A

Brefeldin A (BFA) is a fungal metabolite that is known to inhibit the transport of proteins from the endoplasmic reticulum to the Golgi in *P. falciparum* (Elmendorf and Haldar, 1993). To investigate the timing of PF08_0008 processing, parasites were radiolabelled in the presence of BFA or a solvent control (methanol), and immunoprecipitations were carried out with anti-CTD antibodies as before. The results are shown in figure 6.7.

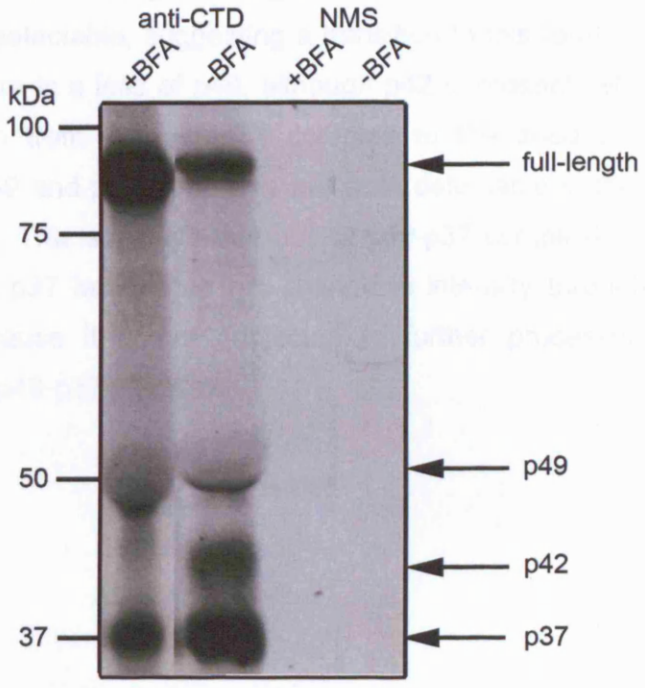
In the presence of methanol, the usual band pattern is observed. This pattern is also observed in the presence of BFA. However, there are marked differences in band intensities between the 2 samples. There appears to be an accumulation of the full length protein in the presence of BFA, relative to the methanol control. There is also much less p37 visible in the presence of BFA, and virtually no p42 products are evident. Any changes in the intensity of p49 cannot be readily determined as the result is obscured by co-migrating IgG heavy chain.

Treatment of parasites with BFA will result in an accumulation of protein between the endoplasmic reticulum and the Golgi. This result indicates that processing of full length PF08_0008 occurs following transport to the Golgi. The p37 processed form of the protein is not completely absent from protein samples labelled in the presence of BFA. This is likely to be due to an incomplete block of protein transport. However, a non-uniform labelling of the protein with ^{35}S-methioine and ^{35}S-cysteine could also be a factor. Parasites are metabolically labelled with Pro-mix, a mix of 70% ^{35}S-methionine and 30% ^{35}S-cysteine. Since 89% of the methionine residues and 69% of the cysteine residues lie in the N-terminal-most 50% of the protein, the N-terminal region will be more heavily ^{35}S-labelled than the C-terminal most 50%.

***Figure 6.7*: Immunoprecipitation of PF08_0008 from brefeldin A-treated parasites**

Autoradiograph showing proteins immunoprecipitated from brefeldin- (+BFA) or methanol- (-BFA) treated schizont material. Anti-CTD antibodies and mouse pre-immune serum (NMS) were used.

The level of processing of PF08_0008 is reduced in the presence of BFA. Full-length PF08_0008 is more abundant in the treated samples than in un-treated samples. The converse is true for the p42 and p37 fragments. It is difficult to determine any difference in the level of p49 in both samples due to co-migrating IgG heavy chain.



6.7 PF08_0008 is processed rapidly following translation

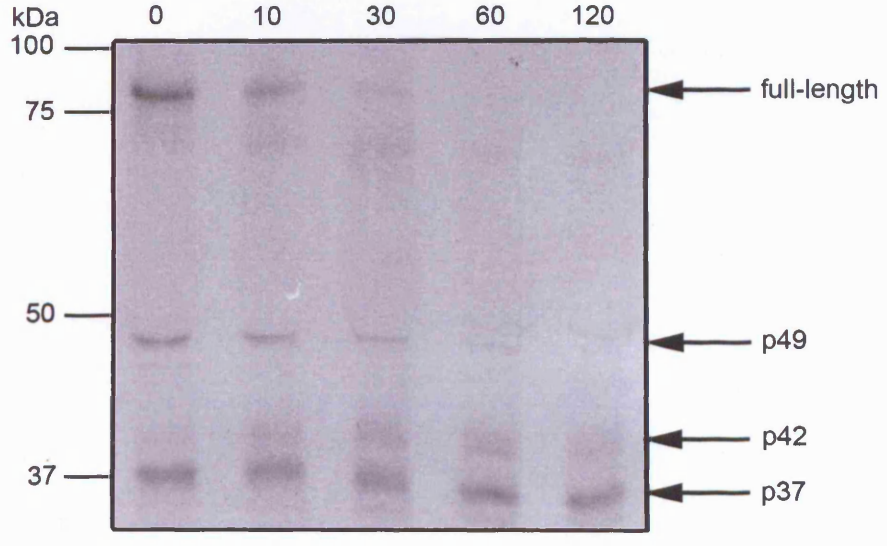
In order to investigate the processing of PF08_0008 further, a pulse-chase experiment was carried out. Tightly synchronised mature schizonts (42 hpi) were metabolically pulse-radiolabelled for 10 minutes with ^{35}S-methionine/cysteine. Non-radiolabelled medium was substituted and parasites were harvested at intervals following the initial labelling. Only proteins translated within this 10 minute window will become radiolabelled, and their fate can be followed throughout the experiment.

The results are shown in figure 6.8. It can be seen that even at t=0, p49 and p37 are already visible. This shows that the processing of full-length protein to give p49-p37 takes place very rapidly, and within 10 minutes of translation. Over time, there is a loss of the full-length PF08_0008. When full-length protein is no longer present, p49 is detectable, suggesting a transition to this form. Towards the end of the chase there is a loss of p49, although p42 is present, which is consistent with a transition from the p49-p37 complex to the shed p42-p37 complex. However, the p49 and p42 fragments are both detectable in the t=10, t=30 and t=60 time points. This suggests that not all p49-p37 complexes are shed at the same time. The p37 band does not change in intensity throughout the chase, presumably because it is not subjected to further processing following the formation of the p49-p37 complex.

***Figure 6.8*: Immunoprecipitation of PF08_0008 from pulse-radiolabelled parasites**

Mature schizonts were pulse-radiolabelled with ^{35}S-methionine/cysteine for 10 minutes and parasites were harvested at 10, 30, 60 and 120 minutes after the end of the 10 minute pulse. A sample of parasites was also taken at the end of the 10 minute pulse. An autoradiograph showing the proteins immunoprecipitated with anti-CTD antibodies at each of these time points is shown.

At t=0, p49 and p37 are already present, suggesting rapid processing of full-length PF08_0008 to p49-p37 occurs, within 10 minutes of translation. Full-length PF08_0008 has disappeared by t=60. Towards the end of the chase there is a loss of p49, although p42 is present which suggests a transition from the p49-p37 complex to the shed p42-p37 complex. p49 and p42 fragments are both detectable in the t=10, t=30 and t=60 time points, suggesting that not all p49-p37 complexes are shed simultaneously. The p37 band does not change in intensity throughout the chase, presumably because it is not subjected to further processing following the formation of the p49-p37 complex.

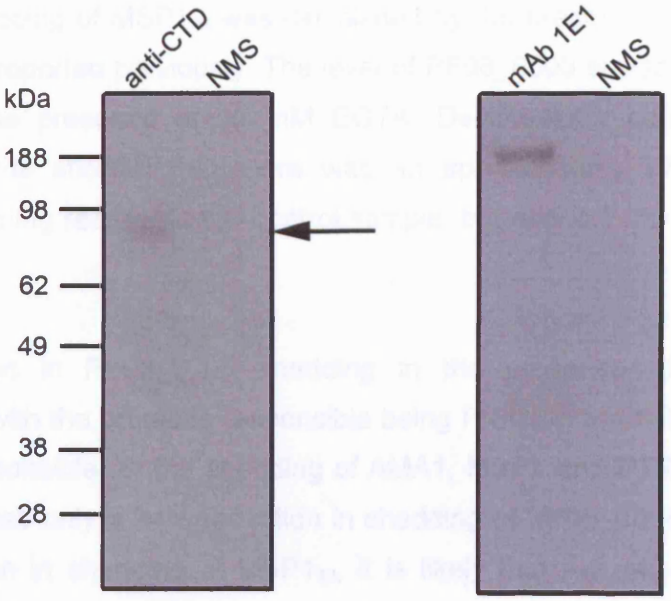


6.8 PF08_0008 is GPI-anchored

In order to determine if PF08_0008 is GPI-anchored as expected, parasites were metabolically labelled with radioactive GPI-anchor precursors. Parasites were labelled with ^{3}H-glucosamine and ^{3}H-mannose and proteins were immunoprecipitated using the PF08_0008 anti-CTD antibodies (figure 6.9). It can be seen from the autoradiograph that ^{3}H-labelled full-length PF08_0008 can be immunoprecipitated with the anti-CTD antibodies, confirming that PF08_0008 is indeed GPI-anchored. The apparent absence of the p50 processing product may be due to its low abundance in the particular preparation used. This result is unlikely to indicate that p50 is no longer GPI-anchored, since figure 6.1b has shown it to be membrane-associated.

***Figure 6.9*: Immunoprecipitation of PF08_0008 from parasites metabolically labelled with the radioactive GPI precursors ^{3}H-glucosamine and ^{3}H-mannose**

a) Anti-PF08_0008 antibodies were used to immunoprecipitate full-length PF08_0008 (arrowed) from labelled parasites. The monoclonal antibody 1E1 was used to immunoprecipitate full-length MSP1 protein (~200 kDa) from labelled parasites. The proteins immunoprecipitated with these two antibodies were separated on different gels and film was subjected to different exposure times. No proteins were immunoprecipitated with normal mouse serum (NMS).



6.9 Shedding of PF08_0008 is largely calcium-independent

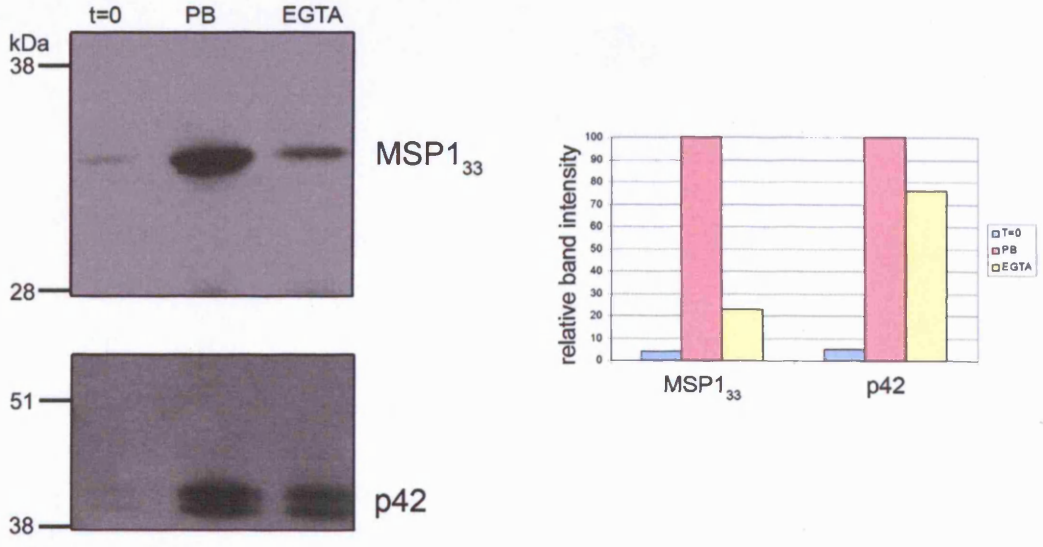
The p42 and p37 fragments of PF08_0008 which can be detected in culture supernatant are likely to be shed by a parasite-derived protease, as is the case with other parasite molecules such as AMA1, PTRAMP, MSP1 and EBA175. In order to investigate the nature of this protease, a merozoite processing assay was carried out. Merozoites were purified in the presence of the calcium-chelating agent EGTA and resuspended in a calcium-containing buffer in order to allow processing to occur. The shedding of $MSP1_{33}$ into supernatant was compared with the shedding of the p42-p37 complex of PF08_0008 (see figure 6.10). It can be seen that the incubation of merozoites at 37°C resulted in an increase in the level of shed $MSP1_{33}$ and p42-p37, relative to the sample that was not incubated at 37°C. Shedding of $MSP1_{33}$ was diminished by the presence of 10 mM EGTA, as has been reported previously. The level of PF08_0008 shedding was reduced slightly by the presence of 10 mM EGTA. Densitometry carried out on the resulting bands showed that there was an approximately 77% reduction in $MSP1_{33}$ shedding relative to the control sample, but only a 24% reduction in p42 shedding.

The reduction in PF08_0008 shedding in the presence of EGTA is not inconsistent with the protease responsible being PfSUB2, the calcium-dependent protease responsible for the shedding of AMA1, MSP1 and PTRAMP. However, since there was only a 24% reduction in shedding of PF08_0008 compared to a 77% reduction in shedding of $MSP1_{33}$, it is likely that the p42-p37 complex is shed by a different protease that is not as calcium-dependent. The rhomboid proteases ROM1 and ROM4 have been shown to be responsible for the shedding of a range of adhesins from the surface of the merozoite in a calcium-independent manner (Baker et al., 2006). However, rhomboids are known to cleave their substrate protein within a transmembrane domain, and PF08_0008 is a GPI-anchored protein. The current evidence suggests that an alternative, probably as yet undefined, parasite protease is responsible for shedding of PF08_0008 into culture supernatant.

***Figure 6.10*: Western blot analysis of the shedding of PF08_0008 from the surface of purified free merozoites**

Merozoites prepared in the presence of EGTA were resuspended in a calcium-containing processing buffer (PB). Parasites were incubated at 37°C in the presence or absence of 10 mM EGTA. Supernatants from merozoites were separated by SDS-PAGE under reducing conditions and probed with anti-$MSP1_{33}$ monoclonal antibody X509 or the anti-CTD antibody. Supernatant from merozoites that had not been subject to the 37°C incubation following re-suspension in processing buffer was also probed, to give an indication of the background level of shed protein fragments present prior to the incubation (at t=0).

It can be see that there was an increase in the amount of shedding of p42 and $MSP1_{33}$ following the incubation in processing buffer, relative to the t=0 control. Addition of EGTA to 10 mM reduced the shedding of $MSP1_{33}$ by 77%. However there was only a 24% reduction in the amount of shed PF08_0008 present. Both values were obtained by performing densitometry on the resultant bands using the ImageJ software. A graph displaying band intensities as percentages relative to the processing buffer controls is shown to the right of the figure.



6.10 Conclusions

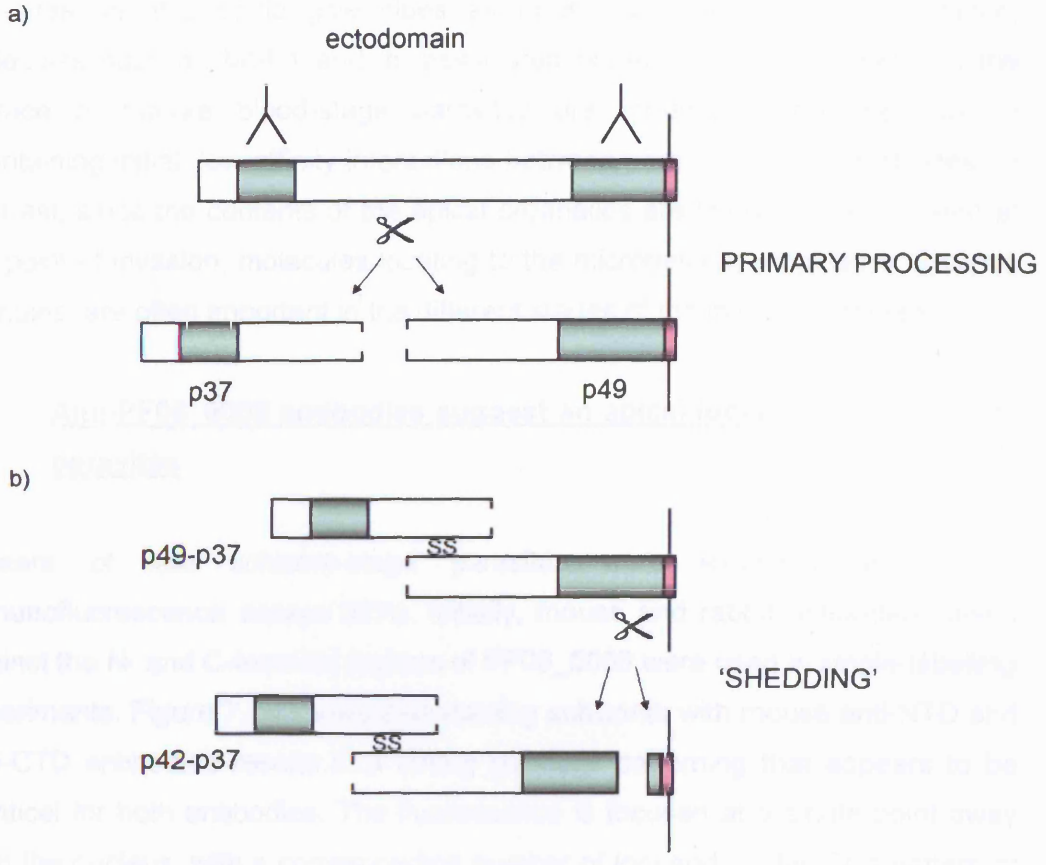
The results that have been presented in this chapter support the following model of proteolytic processing of PF08_0008 (figure 6.11).

PF08_0008 is a GPI-anchored protein, with a predicted molecular mass of 80 kDa. Full length PF08_0008 is proteolytically cleaved, within minutes of translation, to yield two products, a 49 kDa membrane-anchored C-terminal fragment and a 37 kDa N-terminal fragment. Since these fragments are detected in schizonts as well as in free merozoite preparations, this processing must occur prior to schizont rupture. These two fragments remain covalently associated. The p49-p37 complex is shed from merozoites as a result of proteolytic cleavage at two distinct sites near to the membrane anchor. This releases the p42-p37 complex into culture supernatant. The proteases responsible for the primary processing, and also the shedding of PF08_0008, are yet to be identified.

It is likely that either one or both of the p49 and p37 polypeptides run aberrantly on SDS-PAGE gels, since the combined molecular mass of these polypeptides is greater than the predicted molecular mass of the entire protein. This makes it difficult to predict where in the protein the cleavage of full-length PF08_0008 to p49-p37 occurs, and also where the cleavages responsible for the shedding occur.

***Figure 6.11*: Schematic representation of the processing and shedding of PF08_0008**

Primary processing of full length protein (a) gives rise to the p49 and p37 fragments. These fragments are held together in a disulphide-bonded complex and are released into culture supernatant following alternative cleavage events near to the membrane anchor (b), leaving behind a residual 'stub'.



7. PF08_0008 *is located in the micronemes of mature schizonts and is secreted onto the surface of free merozoites*

7.1 Introduction

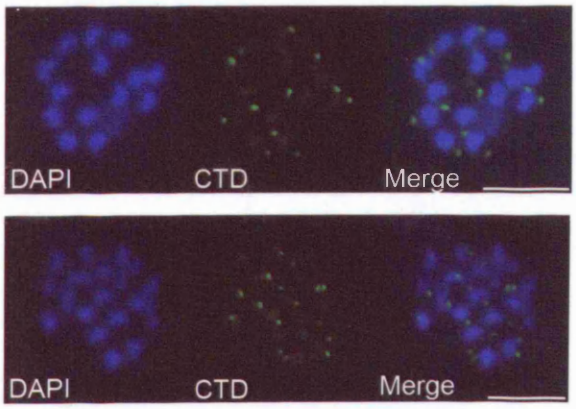
When attempting the characterisation of a novel *P. falciparum* protein such as PF08_0008, it is useful to determine the sub-cellular location of the protein in parasites, as this could give clues as to its functional role. For instance, molecules such as MSP1 and its associated proteins that are present on the surface of mature blood-stage parasites are thought to be important in maintaining initial, low-affinity interactions between parasites and erythrocytes. In contrast, since the contents of the apical organelles are known to be secreted at the point of invasion, molecules locating to the micronemes, rhoptries and dense granules, are often important in the different stages of the invasion process.

7.2 Anti-PF08_0008 antibodies suggest an apical location within mature parasites

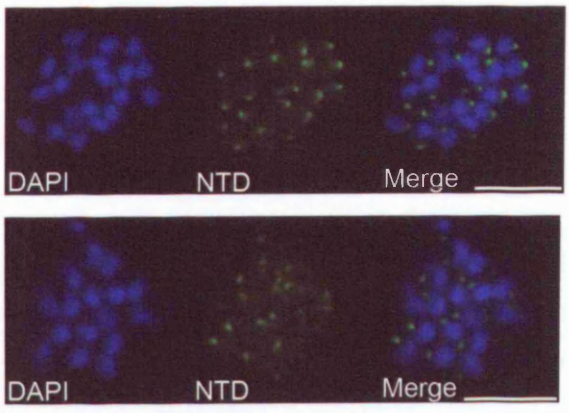
Smears of late schizont-stage parasites were examined in indirect immunofluorescence assays (IFA). Initially, mouse and rabbit antibodies raised against the N- and C-terminal regions of PF08_0008 were used in single-labelling experiments. Figure 7.1a shows that staining schizonts with mouse anti-NTD and anti-CTD antibodies results in a strong punctate patterning that appears to be identical for both antibodies. The fluorescence is focused at a single point away from the nucleus, with a corresponding number of foci and nuclei. This pattern of staining is consistent with a location within the apical organelles. No reactivity is found when schizonts are stained with mouse pre-immune serum (7.1c i).

Figure 7.1b shows that the same results obtained with the mouse antibodies are obtained when the rabbit antibodies are used. The staining with rabbit antibodies seems more dispersed, in contrast to the tight foci obtained with the mouse antibodies. However the overall pattern obtained is similar. This justifies the use of either rabbit or mouse antibodies in future experiments.

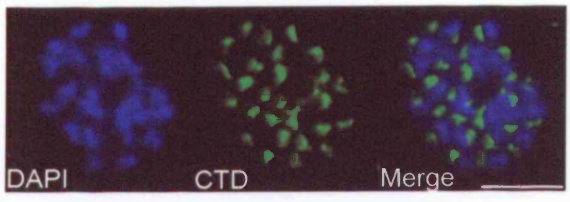
a) i)



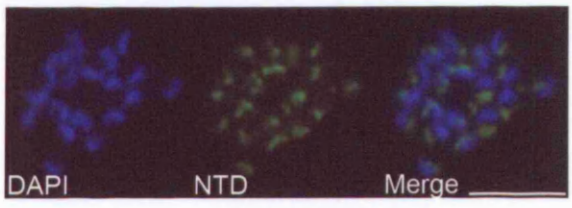
ii)



b) i)

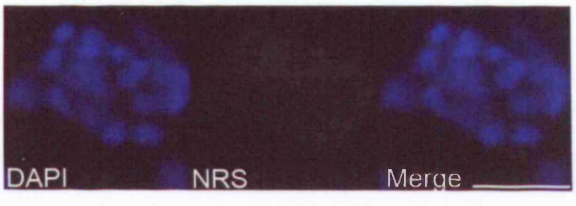


ii)



c) i)

ii)



7.3 PF08_0008 is located within the micronemes of mature *P. falciparum* schizonts

In order to further investigate the location of PF08_0008, dual labelling experiments were carried out using the anti-NTD and anti-CTD antibodies in combination with antibodies raised against other *P. falciparum* proteins with different, defined, sub-cellular locations.

Clag9 encodes a protein that is part of the high-molecular-mass rhoptry protein complex (RhopH) present in the rhoptry bodies of developing merozoites (Ling et al., 2004). RhopH2 is another member of the RhopH complex, and antibodies against both RhopH2 and clag9 were used in dual-labelling experiments.

It can be seen in figure 7.2a that the staining of the C-terminus of PF08_0008 does not colocalise with the staining of either clag9 or RhopH2. Although both patterns are suggestive of an apical location, the red and green staining patterns remain distinct when the two images are overlaid in the final panels. This suggests that PF08_0008, although apically-restricted, is not located in the rhoptries.

Antibodies against two known microneme proteins were also used in dual labelling experiments: AMA1, which is initially present in the micronemes of developing merozoites before being released onto the surface following schizont rupture, and EBA175 which is known to bind the erythrocyte receptor glycophorin A at the tight junction.

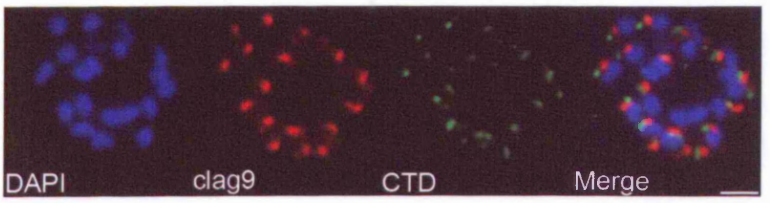
Figure 7.2b shows that anti-PF08_0008 staining colocalises with anti-EBA175 and anti-AMA1 staining in mature schizonts. This is demonstrated by the yellow colour observed when the red and green staining is merged in the far right hand panels. Antibodies against both the N- and C-terminal regions of PF08_0008 give the same result.

***Figure 7.2*: Immunofluorescence assay images showing segmented schizonts dual labelled with anti-PF08_0008 antibodies and antibodies raised against the rhoptry markers RhopH2 and clag9, and the microneme markers EBA175 and AMA1**

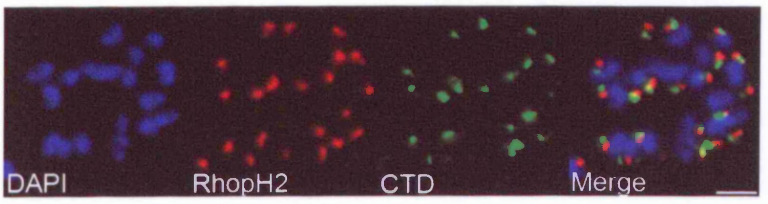
a) Formaldehyde-fixed *P. falciparum* 3D7 schizonts were probed with mouse anti-CTD antibodies and rabbit anti-clag9 antibodies (i) or rabbit anti-RhopH2 antibodies (ii). Rabbit antibodies were detected with an anti-rabbit Texas Red conjugate (red) and mouse anti-CTD antibodies with an anti-mouse FITC conjugate (green). Parasite nuclei were stained with DAPI (blue). Scale bars are all 5 µm. Strong apical FITC and Texas Red signals are exhibited by parasites, although they are distinct and do not overlap.

b) Formaldehyde-fixed *P. falciparum* 3D7 schizonts were probed with mouse anti-CTD or anti-NTD antibodies, and either rabbit anti-EBA175 antibodies (i) or biotinylated anti-AMA1 monoclonal antibody 4G2 (ii). Mouse antibodies were detected with anti-mouse Texas Red (red) or FITC (green) conjugates. Rabbit antibodies were detected with anti-rabbit FITC (green) or Texas Red (red) conjugates. Biotinylated 4G2 was detected with a FITC-streptavidin conjugate (green). Parasite nuclei were stained with DAPI (blue). Scale bars are 2 µm. Strong apical FITC and Texas Red signals are exhibited by parasites, and these overlap to form a yellow colour as visible in the far right hand panels.

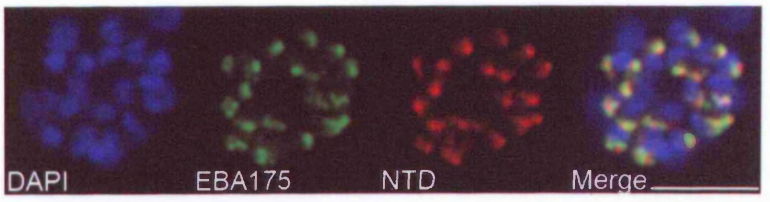
a) i)

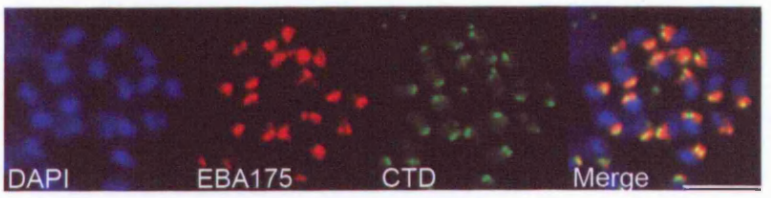

DAPI
clag9
CTD
Merge

ii)

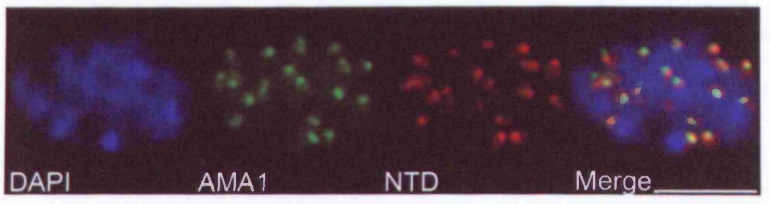

DAPI
RhopH2
CTD
Merge

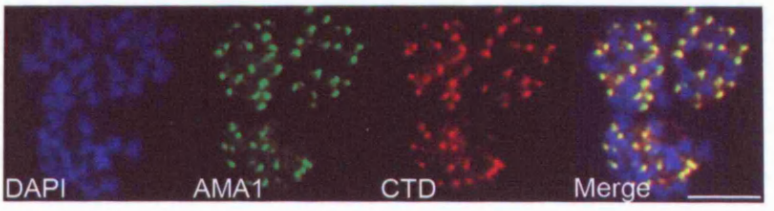
b) i)


DAPI
EBA175
NTD
Merge


DAPI
EBA175
CTD
Merge

ii)


DAPI
AMA1
NTD
Merge


DAPI
AMA1
CTD
Merge

7.4 PF08_0008 is present on the surface of some free merozoites

Some molecules are known to relocate from the apical organelles of segmenting schizonts to the surface of free merozoites, and are subsequently shed at the point of invasion. For example AMA1, which is shed by PfSUB2 (Harris et al., 2005), and the adhesin EBA175, shed by PfROM4 (O'Donnell et al., 2006). In contrast, other apically-restricted molecules do not relocate onto the surface of merozoites, for example RhopH2.

In order to investigate the fate of PF08_0008 following schizont rupture, naturally released merozoites were examined by indirect immunofluorescence in the same way as mature schizonts.

It was evident that PF08_0008 had different distributions in different merozoites, and these are summarised in figure 7.3. In some instances PF08_0008 appeared to be micronemal, as is the case with mature schizonts. Figure 7.3a shows that anti-CTD antibodies give a staining pattern identical to that of anti-AMA1 antibodies, but distinct from the anti-RhopH2 monoclonal antibody 61.3.

Figure 7.3b shows two different distributions of PF08_0008 on free merozoites. In figure 7.3b i, it can be seen that PF08_0008 appears to have a circumferential distribution, similar to that of AMA1. Figure 7.3b ii shows examples of merozoites where PF08_0008 appears to be distributed in a 'cap'-like manner over one pole of the parasite, presumably the apex.

The three different distributions of PF08_0008 described above were not equally represented in the merozoite populations analysed by IFA (see table 7.1).

Table 7.1. The percentage of free merozoites exhibiting one of three observed PF08_0008 distributions is shown as the mean ± one standard deviation for three experiments.

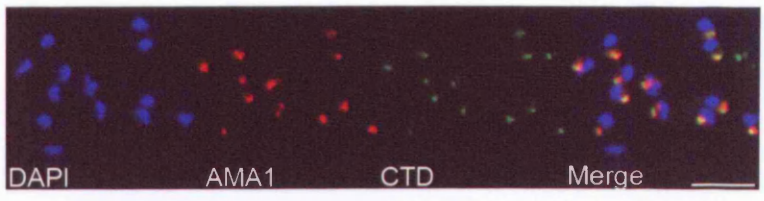
Location	Percent of merozoites
Apical	54±4
Cap	37±7
Surface	11±2

***Figure 7.3*: Immunofluorescence assay images showing free merozoites dual labelled with anti-PF08_0008 antibodies and antibodies raised against the merozoite microneme and surface marker AMA1, and rhoptry marker RhopH2**

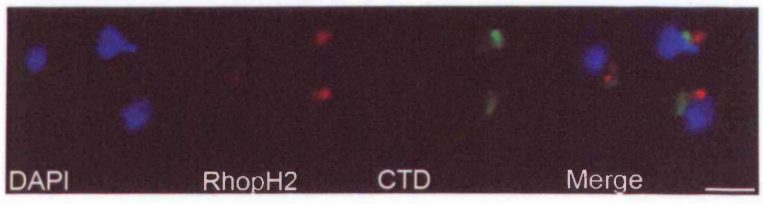
Formaldehyde-fixed *P. falciparum* free merozoites were probed with rabbit anti-CTD antibodies, and either mouse polyclonal anti-AMA1 antibodies (a i and b) or mouse monoclonal anti-RhopH2 antibody 61.3 (a ii). Scale bar of a i is 5 µm, rest are 2 µm. Mouse antibodies were detected with an anti-mouse Alexa Fluor 594 conjugate (red), and rabbit antibodies with an anti-rabbit Alexa Fluor 488 conjugate (green). Parasite nuclei were stained with DAPI (blue).

In 7.3a, there is a clear apical localisation of AMA1, RhopH2 and PF08_0008. AMA1 and PF08_0008 are co-localised in the micronemes of free merozoites. 7.3b i shows a circumferential localisation of both PF08_0008 and AMA1 on the surface of free merozoites. 7.3b ii shows a 'cap'-like distribution of PF08_0008 on the surface of free merozoites (arrowed), in contrast to the circumferential distribution of AMA1.

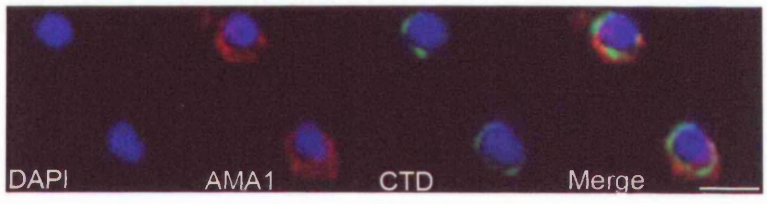
a) i)



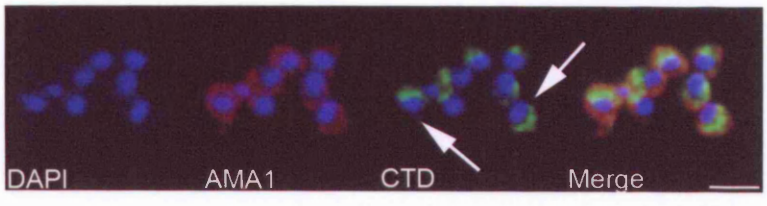
ii)



b) i)



ii)



7.5 Anti-CTD, but not anti-NTD antibodies react with ring-stage parasites

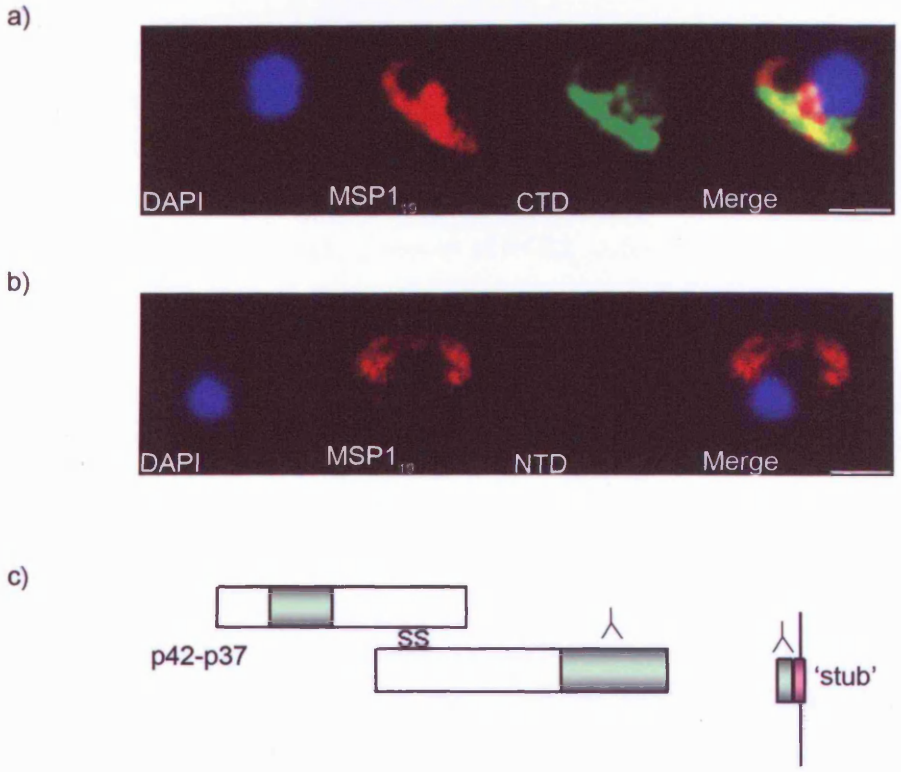
The residual stubs of AMA1 and MSP1, which are generated following the shedding of these molecules from the merozoite surface, are carried into erythrocytes with the invading merozoite, and can be detected by IFA on ring-stage parasites (Blackman et al., 1996; Howell et al., 2005).

IFAs carried out on newly-invaded (≤4 hours post invasion) ring-stage parasites with anti-PF08_0008 antibodies show that the C-terminal part of the protein can be detected in ring-stage parasites whereas the N-terminal part cannot (figure 7.4). Dual labelling IFA shows that the signal produced by the anti-CTD antibodies co-localises with that produced by the monoclonal antibody 1E1, raised against the stub of MSP1, $MSP1_{19}$ (figure 7.6a). In contrast, no signal is produced by the anti-NTD antibodies in IFA on ring-stage parasites (figure 7.4b).

***Figure 7.4*: Immunofluorescence assay images showing newly invaded ring-stage parasites dual labelled with anti-PF08_0008 antibodies and a monoclonal antibody 1E1, raised against $MSP1_{19}$**

Formaldehyde-fixed *P. falciparum* newly-invaded ring-stage parasites were probed with the mouse monoclonal anti-$MSP1_{19}$ antibody 1E1, and either rabbit anti-CTD (a) or anti-NTD (b) antibodies. Mouse antibodies were detected with an anti-mouse Alexa Fluor 594 conjugate (red) and rabbit antibodies with an anti-rabbit Alexa Fluor 488 conjugate (green). Parasite nuclei were stained with DAPI (blue). There was no reactivity exhibited by mouse pre-immune serum with ring-stage parasites (not shown). Scale bars are 2 µm.

This figure shows that there is clearly reactivity of anti-CTD antibodies (a), but not anti-NTD antibodies (b), with ring-stage parasites. Anti-CTD and anti-$MSP1_{19}$ signals appear to co-localise (a), suggesting that 'stub' part of PF08_0008 that is produced upon shedding of the p42-p37 complex is carried into the host cell and remains associated with the plasma membrane of the ring-stage parasite. (c) shows that the specificity of the anti-CTD antibodies (green) for PF08_0008 extends up to the GPI-anchor addition sequence (pink) and that these antibodies would also be reactive with the stub region.



7.6 Conclusions

The results presented above indicate that PF08_0008 is an apically-restricted protein, and is likely to be resident in the micronemes of developing schizonts. This is supported by images demonstrating both the co-localisation of PF08_0008 with two different micronemal markers, AMA1 and EBA175, and the distinction between anti-PF08_0008 labelling and the labelling of rhoptry markers.

It is evident that the antibodies raised against both the N- and C-termini of PF08_0008 show the same pattern of reactivity within schizonts and merozoites. This is not surprising since work detailed in the previous chapter has shown that following a processing event, the N- and C-terminal regions of the protein remain associated through a disulphide bond. This processing event occurs before schizont rupture, and perhaps during transport of PF08_0008 to the micronemes.

Fragments of PF08_0008 are also detectable in culture supernatant, by both the NTD and CTD antibodies. As is the case with MSP1 and AMA1, shedding of PF08_0008 into culture supernatant could generate a residual 'stub' of protein, which remains anchored to the merozoite surface and is carried into the next erythrocyte upon invasion. The IFA data presented here suggests that this is indeed what happens. The specificity of the anti-CTD antibodies for the entire region of PF08_0008 between the putative transmembrane and GPI-anchor sequences means that any 'stub' generated during shedding is also likely to be detected by a subset of these antibodies. This explains the reactivity of the antibodies with ring-stage parasites. In contrast, the N-terminal antibodies do not react with ring-stage parasites, since all of PF08_0008 except the 'stub' would be released into culture supernatant.

Approximately half of the free merozoites examined exhibited an apical distribution of PF08_0008, as is seen in mature schizonts, although smaller proportions demonstrated a 'cap'-like or circumferential distribution. This is not unexpected since the secretion of molecules such as AMA1, EBA175 and SUB2 from the micronemes onto the parasite surface is known to occur in the absence of productive invasion. These data suggest that PF08_0008 is also redistributed from the micronemes of schizonts to the surface of free merozoites. The 'cap'-like distributions could represent an intermediate state, where PF08_0008 has been released onto the surface but has not yet become circumferentially distributed.

The proportion of parasites demonstrating a circumferential or 'cap'-like distribution of PF08_0008 is likely to be directly linked to the proportion of merozoites that are undergoing or have undergone spontaneous microneme secretion in the samples analysed.

It would be of interest to determine how PF08_0008 becomes distributed over the surface of free merozoites. One possibility is that it merely diffuses throughout the plasma membrane following secretion at the apex. Alternatively, PF08_0008 could be linked indirectly, via interactions with other proteins, to the actin-myosin motor of the parasite. This appears to be the case with PfSUB2, as its redistribution towards the rear of the parasite does not occur in the presence of latrunculin A, which causes net actin depolymerisation within the cell (Harris et al., 2005).

PF08_0008 joins the list of previously identified microneme proteins that includes AMA1, EBA140, EBA175, EBA181, MTRAP, ROM4, ROM1, SUB2, PTRAMP, and ASP, with all but the last two having defined roles in the process of invasion. Of these microneme proteins, AMA1, EBA175, SUB2, PTRAMP and ROM4 have also been shown to relocate from the micronemes onto the surface of free merozoites. The evidence presented here is therefore highly suggestive of PF08_0008 having a role in the process of invasion.

8. PF08_0008 *binds erythrocytes but anti-*PF08_0008 *antibodies do not inhibit erythrocyte invasion*

Relatively few *P. falciparum* proteins have been shown to have a well-defined erythrocyte-binding capability. MSP1 is a putative ligand for the erythrocyte protein band 3 (Goel et al., 2003), EBA175 is known to bind glycophorin A on the erythrocyte surface (Sim et al., 1994), EBA140 binds glycophorin C (Maier et al., 2003), and both EBA181 and PfRH1 bind undefined receptors (Gilberger et al., 2003b, Rayner et al., 2001). The presence of PF08_0008 in culture supernatant hints at a potential role for this protein as an adhesin, so this was investigated further.

8.1 The p42-p37 complex of PF08_0008 is able to specifically bind erythrocytes

In order to investigate the ability of PF08_0008 to bind to erythrocytes, supernatant from cultures in which schizonts had undergone rupture was prepared. Samples of culture supernatant were then incubated at 37°C in the presence or absence of washed, type O+ erythrocytes. Figure 8.1a shows that PF08_0008 is present in culture supernatant incubated in the absence of erythrocytes, but not in supernatant incubated in the presence of erythrocytes. The defined erythrocyte adhesin EBA175 is also depleted from culture supernatant following incubation with erythrocytes, whereas the merozoite surface protein MSP3 is not.

Although this result is highly suggestive of PF08_0008 having an erythrocyte-binding ability, it was necessary to establish that the depletion of PF08_0008 from culture supernatant was due to a specific interaction between this protein and erythrocytes. Following incubation of culture supernatant with erythrocytes, the cells were washed and any proteins bound to the surface were eluted with 0.5 M NaCl. A Western blot carried out on the eluted proteins showed that PF08_0008 does indeed bind specifically to the erythrocyte surface (figure 8.1b).

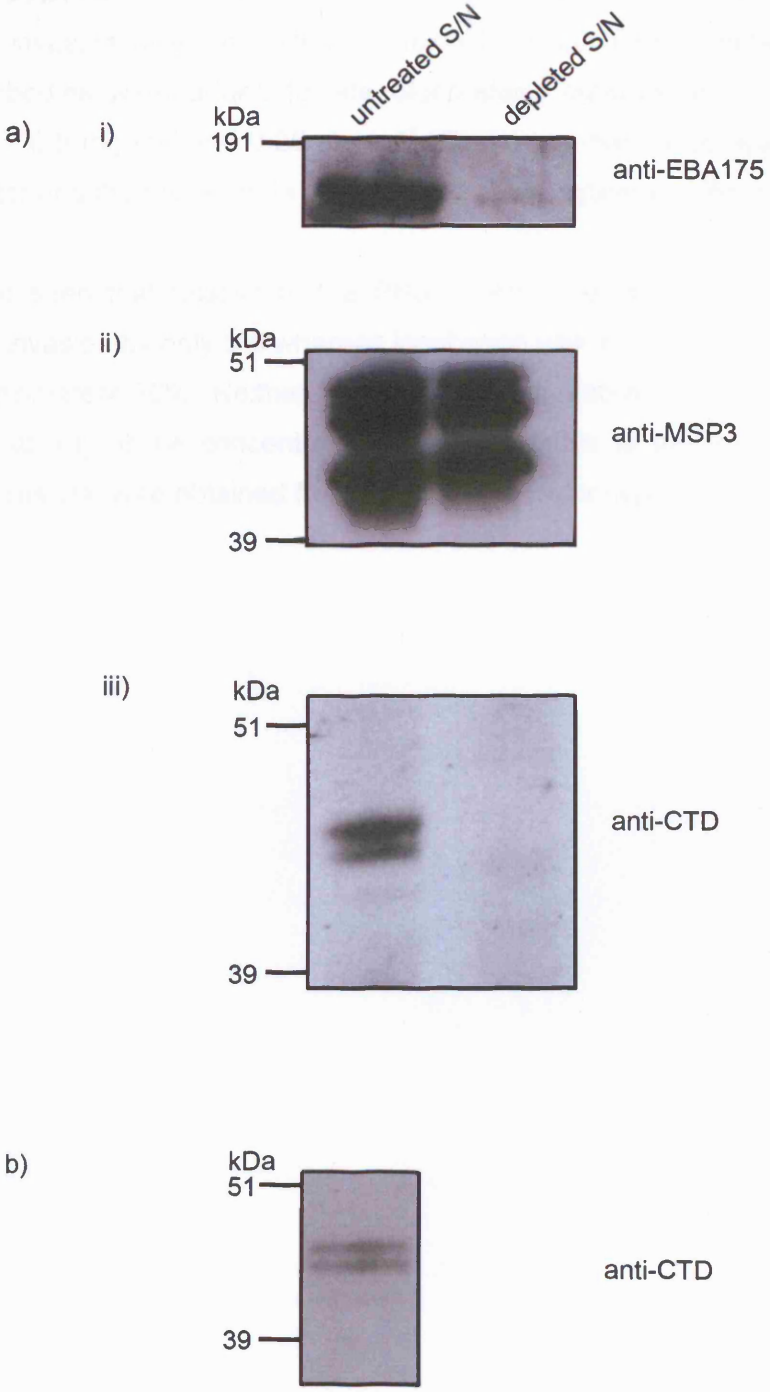
***Figure 8.1*: PF08_0008 binds specifically to the surface of erythrocytes**

a) *P. falciparum* culture supernatant was incubated in the presence or absence of washed, group O+ erythrocytes at 37°C. Cellular material was removed and supernatants were separated by SDS-PAGE under reducing conditions, before being transferred to nitrocellulose and probed with antibodies raised against EBA175 (i), MSP3 (ii) and the C-terminal domain of PF08_0008 (iii). Molecular mass is shown in kDa.

It can be seen that the erythrocyte binding protein EBA175 is present in samples of culture supernatant incubated in the absence of erythrocytes (labelled 'untreated S/N') and is almost completely absent from supernatant incubated in the presence of erythrocytes (labelled 'depleted S/N'), whereas MSP3 is present in both samples. As is the case for EBA175, the C-terminal fragments of PF08_0008 that are released into culture supernatant are depleted following incubation of supernatant with erythrocytes.

b) In order to confirm that the depletion of PF08_0008 from culture supernatant is the result of a specific interaction with erythrocytes, proteins bound to the erythrocyte surface were eluted with 0.5 M NaCl. Eluted proteins were separated by SDS-PAGE under reducing conditions, transferred to nitrocellulose then probed with the anti-CTD antibodies. Molecular mass is shown in kDa.

PF08_0008 can be eluted from the surface of erythrocytes following their incubation with *P. falciparum* culture supernatant, suggesting a specific interaction between this protein and an erythrocyte ligand.


a)
i)
kDa
191
untreated S/N
depleted S/N
anti-EBA175
ii)
kDa
51
39
anti-MSP3
iii)
kDa
51
39
anti-CTD
b)
kDa
51
39
anti-CTD

8.2 Anti-PF08_0008 antibodies do not inhibit erythrocyte invasion

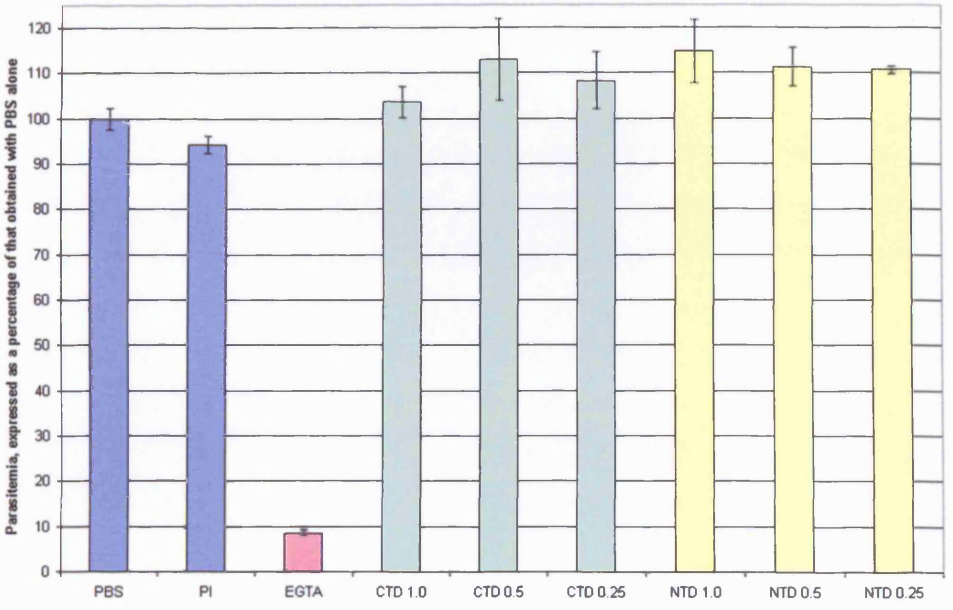
Antibodies raised against EBA175, EBA181 and MSP1, which are all erythrocyte binding proteins, are able to inhibit the invasion of erythrocytes by parasites. In order to address the possibility that the anti-NTD and anti-CTD antibodies inhibit parasite invasion of erythrocytes, invasion inhibition experiments were carried out. Antibodies were added to late blood-stage cultures at concentrations of 1 mg ml^{-1}, 0.5 mg ml^{-1} and 0.25 mg ml^{-1}. Cells were harvested approximately 42 hours later and the parasitemia determined by flow cytometry (figure 8.2).

It can be seen that relative to the PBS control, the rabbit pre-immune serum reduced invasion by only 5% whereas incubation with EGTA reduced parasitemia by approximately 90%. Neither the anti-CTD nor anti-NTD antibodies inhibited invasion at any of the concentrations tested, relative to the control. Essentially identical results were obtained from two identical but independent experiments.

***Figure 8.2*: Anti-PF08_0008 antibodies do not inhibit erythrocyte invasion**

Invasion inhibition assays were carried out using the anti-NTD and anti-CTD antibodies. PBS and rabbit pre-immune serum (labelled PI) were included as negative controls, and 12.5 mM EGTA as a positive control. Rabbit antibodies in PBS were tested at concentrations of 1 mg ml^{-1} (labelled 1.0), 0.5 mg ml^{-1} (labelled 0.5) and 0.25 mg ml^{-1} (labelled 0.25), supplemented with pre-immune IgG as appropriate to give a final IgG concentration of 1 mg ml^{-1} per well.

Parasite cultures were incubated with the test or control solutions, in triplicate, for approximately 42 hours to allow invasion to occur. Cells were stained with ethidium bromide and the parasitemia of each well was determined by flow cytometry. Parasitemia shown is the mean of three wells, relative to the parasitemia resulting from incubation with PBS alone. Error bars indicate one standard deviation above and below the mean. Essentially identical results were obtained in two identical but independent experiments.



8.3 Conclusions

PF08_0008 appears to be an erythrocyte-binding protein - the p42 fragments are depleted from culture supernatant following incubation with erythrocytes, and can be eluted from the erythrocyte surface. However, antibodies raised against the N- and C-terminal regions of PF08_0008 do not inhibit erythrocyte invasion by parasites.

9 *Epitope tagging of endogenous* PF08_0008 *with a C-terminal triple haemagglutinin tag*

It was decided to modify the endogenous copy of *PF08_0008* to include a C-terminal epitope tag, by targeted homologous recombination. The epitope tag of choice was the triple haemagglutinin (HA) tag. The availability of the anti-HA monoclonal antibody 3F10 could in theory allow the purification of tagged PF08_0008 from parasites, with a view to ultimately obtaining information on the cleavage sites responsible for primary processing and shedding. The decision to adopt this strategy was made at a time when the bioinformatic evidence available on PF08_0008 was suggestive of a transmembrane- rather than GPI-anchorage.

9.1 Vector construction

The vector used for transfection is a modified version of the pHH1 vector (Reed et al., 2000), which allows integration of the plasmid into the locus of choice via a single-crossover homologous recombination event (figure 9.2). The vector contains the human dihydrofolate reductase (*dhfr*) gene, for use as a selectable marker.

The 3'-most 829 bp of *PF08_0008* coding sequence was amplified from parasite genomic DNA, incorporating *Bgl*II and *Xho*I restriction sites at the 5' and 3' ends of the product respectively (figure 9.1a). Importantly, primers were designed to omit the last three nucleotides of coding sequence, encoding the stop codon. The PCR product was ligated into the pHH1-T996HA3 vector (a gift from Pippa Harris) (Harris et al., 2005), which was obtained in a *Bgl*II- and *Xho*I-digested form, to make pHH1-PF08_0008HA3, with a continuous reading frame formed from the *PF08_0008* coding region into the triple HA tag coding sequence.

E. coli TOP10 cells were transformed with pHH1-PF08_0008HA3 and a PCR screen was carried out on a number of resulting colonies (figure 9.1b), using the same primers as above. Plasmid DNA was obtained from a number of colonies that appeared to contain pHH1-PF08_0008HA3 and a restriction digest was carried out to confirm that this was indeed the case (figure 9.1c). The plasmid DNA from a number of clones was also sequenced to ensure no mistakes had been introduced during PCR amplification of the *PF08_0008* coding region.

***Figure 9.1*: Cloning the 3'-most 829 bp of the *PF08_0008* coding sequence into pHH1-T996HA3**

a) Primers were designed to amplify the 3'-most 829 bp of the *PF08_0008* coding sequence, excluding the final three nucleotides that encode the stop codon. *Xho*I and *Bgl*II sites were incorporated into the 5' and 3' ends of the PCR product respectively. A product of 848 bp was expected. Analysis of PCR products was carried out by agarose gel electrophoresis. Molecular sizes are given in bp.

b) A PCR screen was carried out on *E. coli* colonies obtained following transformation with pHH1-PF08_0008HA3. Products were analysed by agarose gel electrophoresis. Molecular sizes are given in bp. Of the three colonies shown (numbered 1-3), colonies 1 and 2 appear to contain the vector with the insert.

c) Restriction digests were carried out on plasmid DNA obtained from the two colonies that gave a positive PCR screen result in b. DNA was digested with *Xho*I and *Bgl*II in order to release an 841 bp fragment from the vector (arrowed). The size of the remaining plasmid backbone is 5867 bp. Products were analysed by agarose gel electrophoresis. Molecular sizes are given in bp.

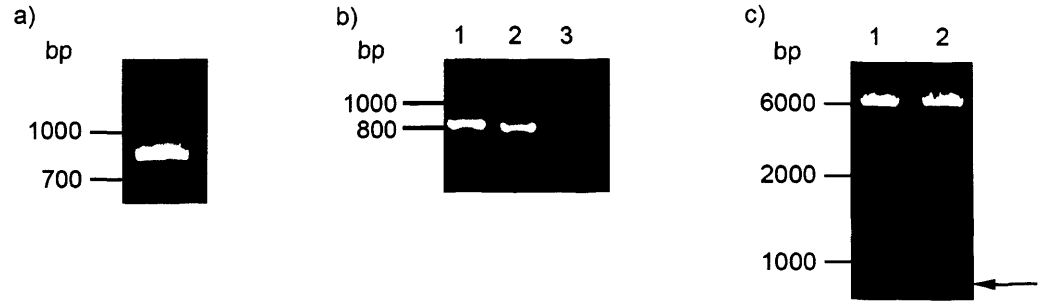


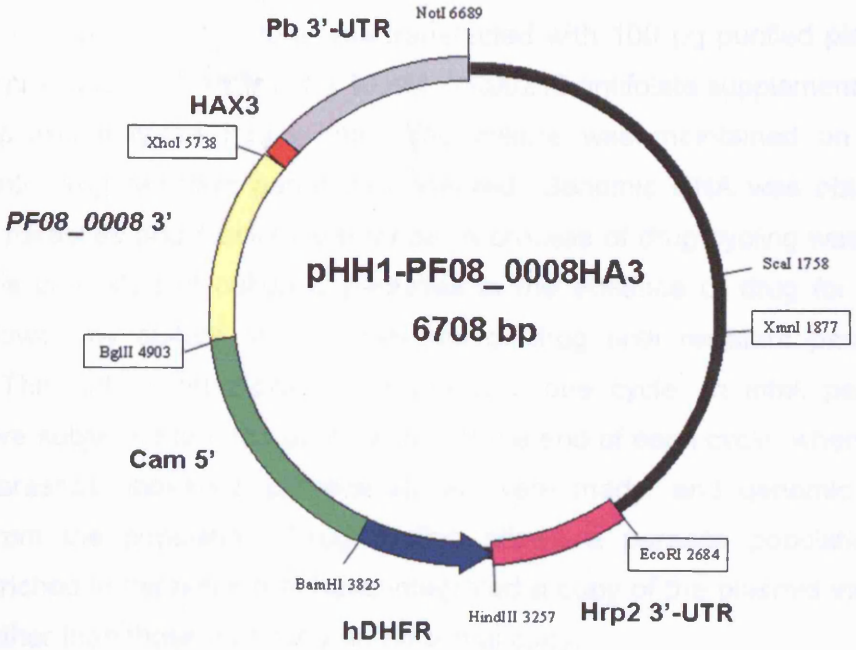
***Figure 9.2*: The vector pHH1-PF08_0008HA3**

a) The vector used to transfect parasites, pHH1-PF08_0008HA3, consists of the 3'-most 829 bp of the *PF08_0008* coding sequence, excluding the terminal stop codon (yellow), fused in-frame to the coding sequence of the triple haemagglutinin tag (red). The correct transcription termination and polyadenlyation of this gene is regulated by the presence of the *P. berghei* dihydrofolate reductase 3' untranslated region (UTR) (grey). The human dihydrofolate reductase (hDHFR) cassette (blue) is under the transcriptional control of the 5' untranslated region of the *P. falciparum* calmodulin gene (green) and the 3' untranslated region of the *hrp*2 gene (purple). The *Xho*I and *Bgl*II sites used to clone in the *PF08_0008* fragment are boxed, as are the *Xmn*I and *Eco*RI sites that were used to generate the probe for Southern blot analysis. The *Sca*I site in the plasmid backbone is also indicated.

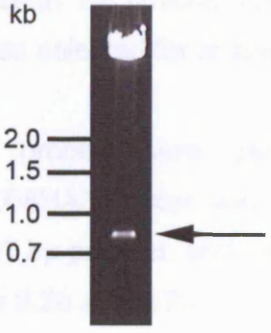
b) The pHH1-PF08_0008HA3 vector was digested with *Xmn*I and *Eco*RI in order to generate an 807 bp fragment for use as a probe in a Southern blot analysis. The products of the digest were analysed by agarose gel electrophoresis. The size of the remaining plasmid backbone is 5901 bp. The probe (arrowed) was gel purified and again analysed by agarose gel electrophoresis (ii). Molecular sizes are given in kb.

c) The pHH1-PF08_0008HA3 vector was digested with *Xho*I and *Bgl*II in order to generate an 841 bp fragment for use as a probe in Southern blot analysis. The products of the digest were analysed by agarose gel electrophoresis. The size of the remaining plasmid backbone is 5867 bp. The probe (arrowed) was gel purified and again analysed by agarose gel electrophoresis (ii). Molecular sizes are given in kb.

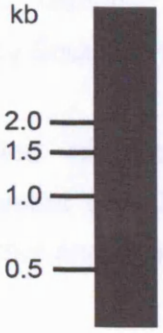
a)


Pb 3'-UTR
NotI 6689
HAX3
XhoI 5738
PF08_0008 3'
pHH1-PF08_0008HA3
6708 bp
ScaI 1758
XmnI 1877
BglII 4903
Cam 5'
EcoRI 2684
BamHI 3825
HindII 3257
hDHFR
Hrp2 3'-UTR

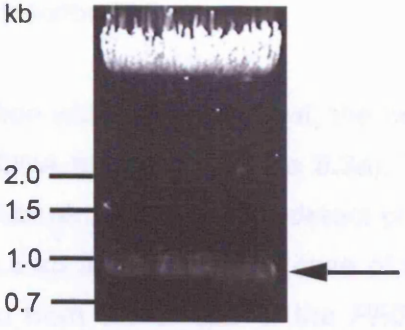
b) i)


kb
2.0
1.5
1.0
0.7

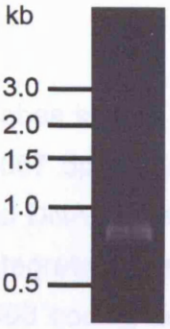
ii)


kb
2.0
1.5
1.0
0.5

c) i)


kb
2.0
1.5
1.0
0.7

ii)


kb
3.0
2.0
1.5
1.0
0.5

9.2 Transfection of parasites

A ring stage *P. falciparum* culture was transfected with 100 µg purified plasmid DNA. Drug pressure, in the form of a 10 nM WR99210 antifolate supplement, was applied approximately 24 hours later. The culture was maintained on drug pressure until drug resistant parasites appeared. Genomic DNA was obtained from these parasites and stocks were made. A process of drug cycling was then started. This consisted of culturing parasites in the absence of drug for three weeks, followed by culture in the presence of drug until resistant parasites appeared. This 'off/on' drug pressure represents one cycle. In total, parasite cultures were subjected to three drug cycles. At the end of each cycle, when drug resistant parasites appeared, parasite stocks were made, and genomic DNA obtained from the population. Drug cycling allows a parasite population to become enriched in parasites that have integrated a copy of the plasmid into the genome, rather than those that carry an episomal copy.

9.3 Southern blot analysis

Following three rounds of drug cycling, individual clones were obtained from the parasite population by limiting dilution. Genomic DNA was obtained from a number of clones selected for analysis by Southern blot.

Two different probes were generated for Southern blot analysis. The pHH1-PF08_0008HA3 vector was digested with *Xmn*I and *Eco*RI in order to generate an 807 bp product, and with *Xho*I and *Bgl*II in order to excise an 841 bp product (figures 9.2b and 9.2c).

Genomic DNA from *P. falciparum* 3D7 parasites, along with DNA from parasite populations obtained at the end of each drug cycle, and from individual clones, was digested with *Sca*I and *Swa*I and subjected to Southern blot analysis using the two probes described above.

Following digestion with *Sca*I and *Swa*I, the two probes should detect a different combination of DNA fragments (figure 9.3a). The 807 bp probe is derived from the plasmid backbone, and so would detect plasmid DNA in the form of a 6.7 kb episome, or a 5.8 kb fragment in the case of an integrated plasmid. The 841 bp probe is derived from the 3' end of the *PF08_0008* coding sequence, and so

would detect a 3.8 kb fragment in the case of an unmodified *PF08_0008* locus, or 4.8 kb and 5.8 kb bands in the case of an integrated plasmid.

The results of the Southern blot are shown in figure 9.3b and it can be seen that in the case of four out of five parasite clones analysed, there appears to have been integration of the pHH1-PF08_0008HA3 plasmid into the endogenous *PF08_0008* locus, with no remaining episome evident. Clone B4 was chosen for use in future experiments. The integration event was confirmed by sequencing a PCR product amplified from clone B4 genomic DNA using a forward primer from outside the region of *PF08_0008* that was cloned into the pHH1-T996HA3 vector, and a reverse primer from the plasmid, downstream of the triple HA tag coding sequence. This confirmed that the sequence of *PF08_0008* was in-frame with the triple HA tag coding sequence.

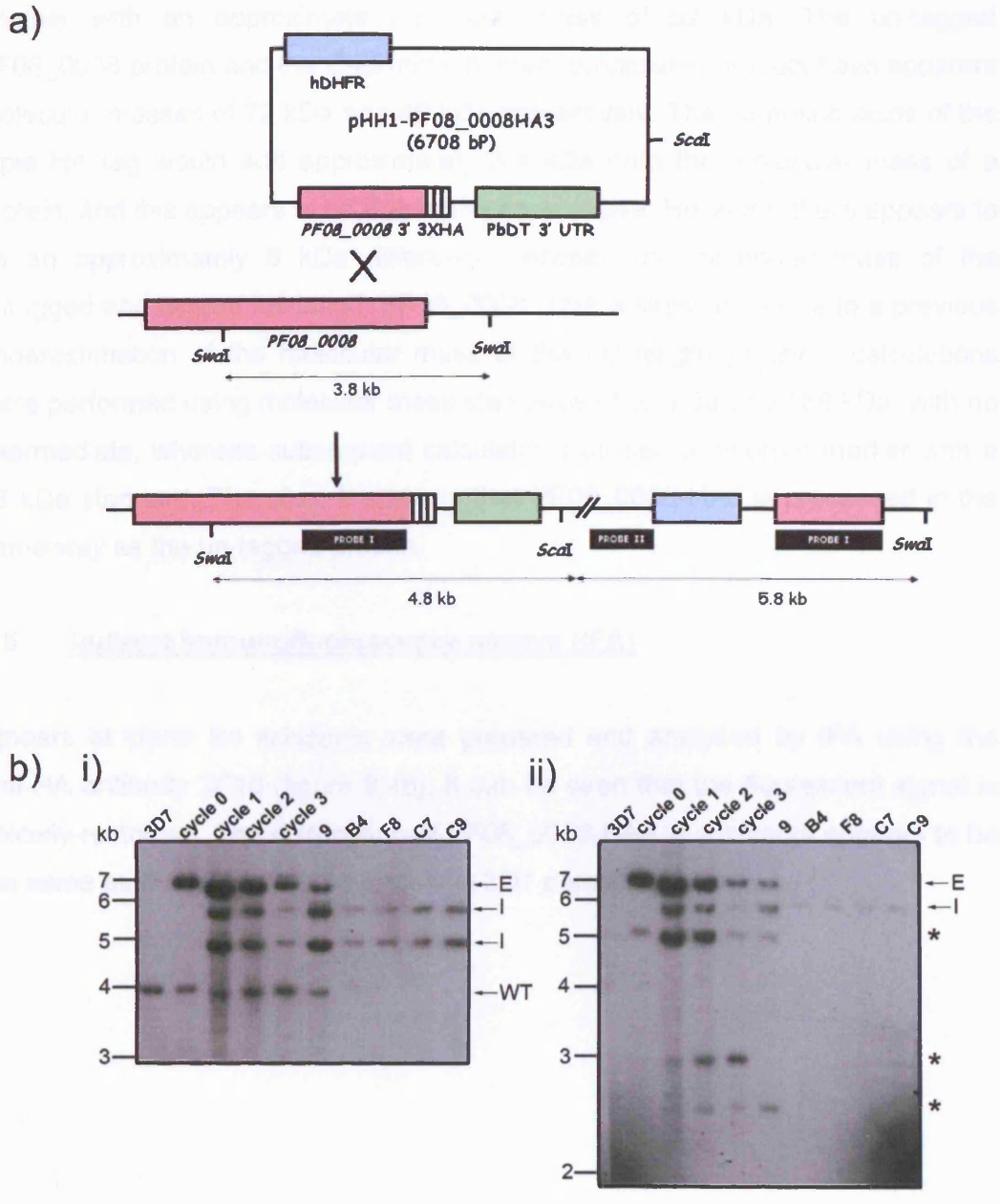
***Figure 9.3*: Southern blot analysis of parasites transfected with pHH1-PF08_0008HA3**

a) Schematic representation of the integration plasmid pHH1-PF08_0008HA3 and a single-crossover homologous recombination event. Integration reconstitutes a full-length *PF08_0008* gene with a triple HA tag at the 3' end. The correct transcription termination and polyadenlyation of this gene is regulated by the presence of the *P. berghei* dihydrofolate reductase 3' untranslated region (UTR) (PbDT 3' UTR). The *Sca*I and *Swa*I sites are indicated, as well as the size of the fragments obtained following a double restriction digest with these enzymes, in the case of an unmodified *PF08_0008* locus (3.8 kb), an episome (6.7 kb) and a single-crossover integration event (4.8 kb and 5.8 kb). The position of the two probes used is indicated: probe I (produced by the *Xho*I/*Bgl*II digest of pHH1-PF08_0008HA3) will recognise the 3.8kb, 4.8 kb and 5.8 kb fragments, and probe II (produced by the *Eco*RI/*Xmn*I digest of pHH1-PF08_0008HA3) will recognise the 5.8 kb and 6.7 kb fragments.

b) Southern blots of *Sca*I/*Swa*I-digested genomic DNA from: wild-type 3D7 parasites (3D7), the first population of WR92210-resistant parasites that appeared following transfection (cycle 0), the populations of resistant parasites present after one, two or three rounds of drug cycling (cycle 1, cycle 2 and cycle 3) and five clones derived by limiting dilution (9, B4, F8, C7 and C9). Blots were probed with i) probe I or ii) probe II. Molecular sizes are indicated in kb.

i) 3D7 parasites show reactivity with probe I, with a band at approximately 3.8 kb, consistent with an unmodified *PF08_0008* locus (labelled WT). Cycles 0-3 and clone 9 show evidence of episome in the form of a 6.7 kb band (E), and also an unmodified *PF08_0008* locus. Cycles 1-3 and clone 9 also appear to contain parasites with an integrated plasmid, represented by the two bands at 4.8 kb and 5.8kb (I), and are thus mixed populations at this stage. Only clones B4, F8, C7 and C9 demonstrate an absence of an unmodified *PF08_0008* locus, confirming that they are true clonal populations. Episome is also absent from these 4 clones.

ii) 3D7 parasites do not react with probe II, as expected since the probe is derived from the plasmid backbone. Episome is present in cycles 0-3 (E), and also in clone 9. Cycles 1-3 and clone 9 also appear to contain parasites with an integrated plasmid, represented by the band at 5.8 kb (I), and are thus mixed populations. Clones B4, F8, C7 and C9 show evidence of an integrated plasmid in the absence of episome. In parasites from cycles 1-3, and in clone 9, there are also additional unexpected bands present of sizes 5 kb, 3 kb and 2.5 kb (labelled *). The presence of these bands cannot be explained by an event such as concatemerisation of the transfection plasmid and are of unknown origin. Importantly, they are absent from the clonal parasite populations.


a)
hDHFR
pHH1-PF08_0008HA3
(6708 bP)
ScaI
PF08_0008 3' 3XHA
PbDT 3' UTR
PF08_0008
SwaI
SwaI
3.8 kb
PROBE I
PROBE II
PROBE I
SwaI
ScaI
SwaI
4.8 kb
5.8 kb
b) i)
kb
3D7
cycle 0
cycle 1
cycle 2
cycle 3
9
B4
F8
C7
C9
7
6
5
4
3
E
I
I
WT
ii)
kb
3D7
cycle 0
cycle 1
cycle 2
cycle 3
9
B4
F8
C7
C9
7
6
5
4
3
2
E
I
*
*
*

9.4 Western blot analysis

Western blot analysis was carried out on clone B4 schizont material (figure 9.4a) using the anti-HA antibody 3F10. The results are consistent with the processing model that was discussed in chapter 6. This antibody recognises two polypeptides fragments: one with an approximate molecular mass of 80 kDa and another with an approximate molecular mass of 52 kDa. The un-tagged PF08_0008 protein and the C-terminal primary processing product have apparent molecular masses of 72 kDa and 49 kDa respectively. The 33 amino acids of the triple HA tag would add approximately 3-4 kDa onto the molecular mass of a protein, and this appears to be true in the case of p49. However, there appears to be an approximately 8 kDa difference between the molecular mass of the untagged and tagged full-length PF08_0008. This is likely to be due to a previous underestimation of the molecular mass of the full length protein - calculations were performed using molecular mass standards of 62 kDa and 188 kDa, with no intermediate, whereas subsequent calculations utilised a different marker with a 98 kDa standard. Therefore it appears that PF08_0008-HA3 is processed in the same way as the un-tagged protein.

9.5 Indirect immunofluorescence assays (IFA)

Smears of clone B4 schizonts were prepared and analysed by IFA using the anti-HA antibody 3F10 (figure 9.4b). It can be seen that the fluorescent signal is apically-restricted. The distribution of PF08_0008-HA3 in schizonts appears to be the same as that of un-tagged protein in 3D7 parasites.

***Figure 9.4*: Western blot and indirect immunofluorescence analyses of transgenic parasites expressing PF08_0008-HA3**

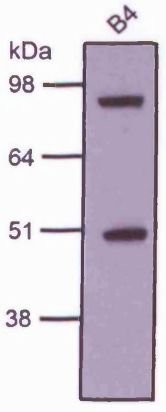
a)

i) Clone B4 schizont samples were lysed directly into reducing sample buffer, separated by SDS-PAGE, transferred to nitrocellulose and probed with the anti-HA monoclonal antibody 3F10. Molecular mass is indicated in kDa.

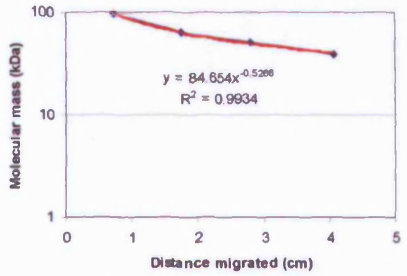
ii) The distances travelled along the gel by the protein standards were measured and used to plot a standard curve of distance against $\log_{10}$(molecular mass). Data points are connected in black, and the trendline displayed in red. This enabled the calculation of the molecular mass of the two polypeptides shown in i), using their migration distances and the equation displayed. The R^2 value is also indicated on the graph. The table shows the data used to produce the standard curve. The values marked with an asterisk represent data for the two polypeptides.

b) Smears of clone B4 schizonts were fixed with formaldehyde and probed with the rat anti-HA monoclonal antibody 3F10. 3F10 was detected with an anti-rat Alexa Fluor 488 conjugate. Parasite nuclei were stained with DAPI (blue). Scale bar indicates 5 µm.

a) i)

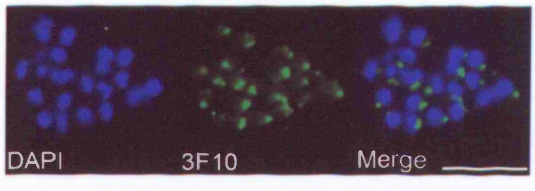


ii)



Distance migrated (cm)	Molecular mass (kDa)
0.75	97
1.75	64
2.8	51
4.075	39
1.05*	80
2.46*	52

b)



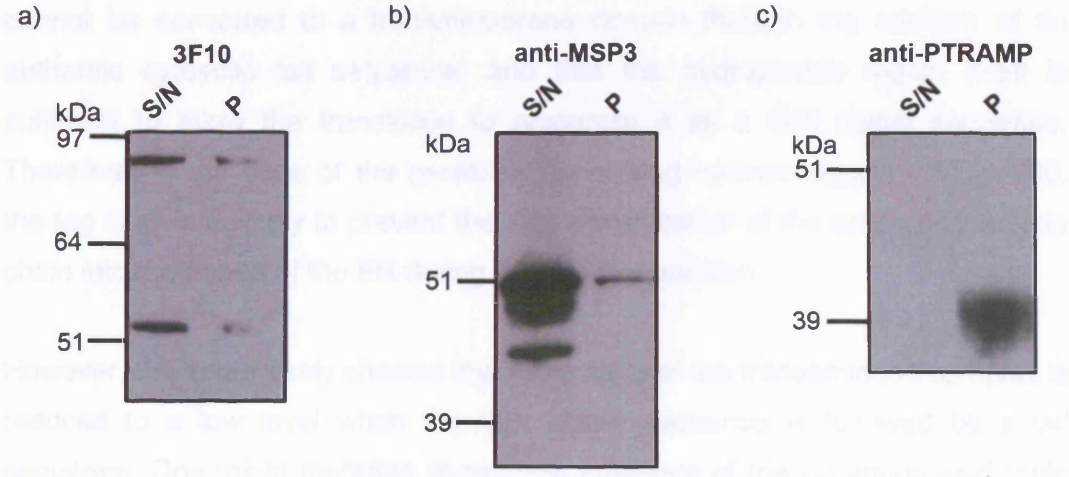
9.6 Solubility of PF08_0008-HA3

In order to determine whether epitope-tagged PF08_0008 is still membrane-associated, like the wild-type protein, clone B4 schizont material was treated with a high *p*H carbonate buffer, and the soluble and insoluble fractions were analysed by Western blot with the anti-HA antibody 3F10 (figure 9.5). It is evident that the vast majority of epitope-tagged PF08_0008 is no longer associated with membrane, as it is present in the carbonate-soluble fraction. The low level of protein detected in the carbonate-insoluble fraction could either be due to an incomplete solubilisation of proteins with the high *p*H carbonate buffer, or the presence of a proportion of epitope tagged protein that is membrane anchored. The detection of a low level of MSP3, which is a peripheral merozoite surface protein, in the carbonate-insoluble fraction suggests that the former is likely.

***Figure 9.5*: Separation of membrane-associated and soluble proteins from transgenic parasites expressing PF08_0008-HA3**

Clone B4 schizont preparations were treated with a high *p*H carbonate buffer (100 mM Na_2CO_3 buffer (*p*H 11.0)). Fractions of protein soluble (S/N) and insoluble (P) in this buffer were obtained, and were separated under reducing conditions by SDS-PAGE. Proteins were transferred to nitrocellulose, and probed with the anti-HA monoclonal antibody 3F10 (a), as well as antibodies raised against the soluble merozoite surface protein MSP3 (b), and the transmembrane protein PTRAMP (c). Molecular masses are given in kDa.

The majority of PF08_0008-HA3 appears in the carbonate-soluble fraction, indicating a lack of membrane association. A small proportion is present in the carbonate-insoluble fraction, which could be an indication of incomplete solubilisation, or a subset of PF08_0008-HA3 that is membrane-anchored. MSP3 appears almost entirely in the carbonate-soluble fraction, as expected. The presence of a small amount in the carbonate-insoluble fraction is likely to be due to an incomplete solubilisation of protein in the carbonate buffer. PTRAMP appears exclusively in the carbonate insoluble fraction, indicating a membrane association.



9.7 Conclusions

Transgenic parasites expressing an epitope tagged version of PF08_0008 have been generated through targeted homologous recombination. None of the clones exhibited any growth defect relative to the parental 3D7 clone, suggesting that neither the fusion of the triple haemagglutinin tag to the C-terminus of PF08_0008, nor the replacement of the endogenous *PF08_0008* 3' UTR with the heterologous *P. berghei* sequence, had any detrimental effect on parasite replication.

One notable difference from wild-type parasites is the solubility of PF08_0008-HA3 in high *p*H carbonate buffer. This suggests that the protein is no longer membrane anchored in the transgenic parasites. This is consistent with a GPI- rather than transmembrane-anchorage of PF08_0008 in 3D7 parasites. As discussed earlier, GPI-anchor addition begins with the translocation of the entire polypeptide chain into the lumen of the endoplasmic reticulum, where the GPI signal sequence is recognised by a transamidase complex. Even when the hydrophobic region of the GPI signal sequence is followed by an authentic charged cytosolic tail sequence, as is the case with transmembrane signal sequences, the entire polypeptide chain is still translocated into the lumen of the ER (Dalley and Bulleid, 2003a). This indicates that a GPI-anchor signal sequence cannot be converted to a transmembrane domain through the addition of an authentic cytosolic tail sequence, and that the hydrophobic region itself is sufficient to allow the translocon to recognise it as a GPI signal sequence. Therefore, in the case of the parasites expressing epitope-tagged PF08_0008, the tag itself is unlikely to prevent the total translocation of the entire polypeptide chain into the lumen of the ER during GPI-anchor addition.

However, the same study showed that the activity of the transamidase complex is reduced to a low level when the GPI signal sequence is followed by a tail sequence. One might therefore expect the presence of the 33 amino acid triple HA tag sequence C-terminal of the GPI-anchor addition sequence of PF08_0008 to impair the ability of the transamidase complex to efficiently recognise the GPI signal sequence, resulting in PF08_0008 that is mostly soluble in its epitope-tagged form.

The tagged form of the protein appears to be processed in the same way as the un-tagged form. The full length protein is detectable by Western blot in schizont extracts, as well as what is likely to be a tagged version of the p49 primary processing product. These results are also consistent with the two-step processing described in chapter 6.9. 3F10 would not be expected to detect the N-terminal p37 fragment due to the location of the triple HA epitope at the extreme C-terminus of the protein, and for the same reason one would not expect the p42 shed fragments to be detected either.

PF08_0008-HA3 also appears to be apically restricted, like its non-tagged counterpart, although immuno-electron microscopy would be necessary to confirm the specific sub-cellular location. These results suggest that PF08_0008 need not be membrane anchored in order to be trafficked correctly or proteolytically processed. Moreover, the presence of a membrane-bound form of PF08_0008 is not essential for parasite growth, or clones would not have been established.

10. Discussion

The data that have been presented here suggest that PF08_0008 is a GPI-anchored microneme protein that is proteolytically processed within schizonts and ultimately shed from the surface of merozoites. Culture supernatant depletion experiments have also shown that PF08_0008 binds specifically to the surface of erythrocytes.

PF08_0008 is the twelfth GPI-anchored protein to be identified in the asexual blood stage of *P. falciparum*, and the second to be found to be micronemal following immunofluorescence analysis. Nine out of the twelve previously identified GPI-anchored proteins appear to be surface-resident proteins. PF08_0008, RAMA and ASP may be part of a subset of GPI-anchored proteins that are initially present in the apical organelles rather than on the merozoite surface.

It is likely that PF08_0008 is a micronemal protein. Micronemes discharge their contents following the reorientation of the parasite. However, it is thought that low levels of protein are released from apical organelles onto the surface of the parasite prior to reorientation and the formation of the tight junction, and may contribute to initial attachment. PF08_0008 is also shed from the surface of the parasite into culture supernatant and these two observations are suggestive of PF08_0008 functioning either in the initial attachment of the merozoite to the erythrocyte, and/or at some point between the formation of the tight junction and the completion of invasion, at which point the protein is presumably shed from the surface of the parasite.

Anti-PF08_0008 antibodies do not inhibit erythrocyte invasion in the concentration range tested. This does not appear to be consistent with PF08_0008 playing a role in invasion; antibodies against EBA175 and EBA181, which are both known erythrocyte-binding proteins, MSP1, which is presumed to be an erythrocyte-binding protein, and AMA1, which is critical in invasion, have all been shown to inhibit invasion. It could be argued, however, that antibodies raised against PF08_0008 would only inhibit invasion if they directly or indirectly interfered with a critical physical interaction necessary for the functioning of the protein. For example a receptor-ligand interaction could be disrupted, or a

processing site blocked. However, it is entirely possible that regions of the protein important in maintaining the functionality of the protein lie outside of the regions against which antibodies were raised. Indeed, less than one third of the PF08_0008 sequence was included in the recombinant proteins against which antibodies were raised. Additionally, PF08_0008 may not be sufficiently exposed to antibody during the invasion process, if for example, it is concentrated at the moving junction, which would be largely inaccessible to antibody.

Moreover, the anti-NTD and anti-CTD antibodies that have been generated during this project were both raised against insoluble, denatured protein. One might expect that such antibodies do not detect native parasite PF08_0008 in its correctly folded state. Although IFA has shown that this is certainly not the case, with these antibodies specifically detecting an apically-located parasite protein, it is possible that the binding of these antibodies to their epitope might be impaired relative to any antibodies raised against correctly folded protein. The anti-NTD and anti-CTD antibodies may be able to detect native parasite protein by IFA, but may not be of sufficiently high avidity to disrupt the functioning of the protein.

Since PF08_0008 binds specifically to erythrocytes, it could be acting as an adhesin that is connected to the motor complex, albeit indirectly due to its GPI-anchorage, and allows productive invasion to occur, in the same manner as proteins such as EBA175.

It has been demonstrated in *P. knowlesi* that AMA1 is essential for the parasite to reorientate, possibly via a concentration gradient of adhesive interactions, and it is likely to be distributed on the surface of the merozoite before the formation of the tight junction (Mitchell et al., 2004). The vast majority of free merozoites analysed with anti-PF08_0008 antibodies displayed a punctuate or 'cap'-like apical staining pattern, with few displaying any circumferential distribution. This is consistent with a model in which PF08_0008 is released upon formation of the tight junction with the erythrocyte, rather than before initial attachment.

The moving junction is believed to function to exclude host cell membrane proteins from entering the nascent parasitophorous vacuole membrane (PVM) (Mordue et al., 1999), as well as being the site of removal of merozoite surface molecules such as MSP1 during invasion. In *T. gondii*, the moving junction complex is minimally composed of TgAMA1 and TgRON4 (Alexander et al.,

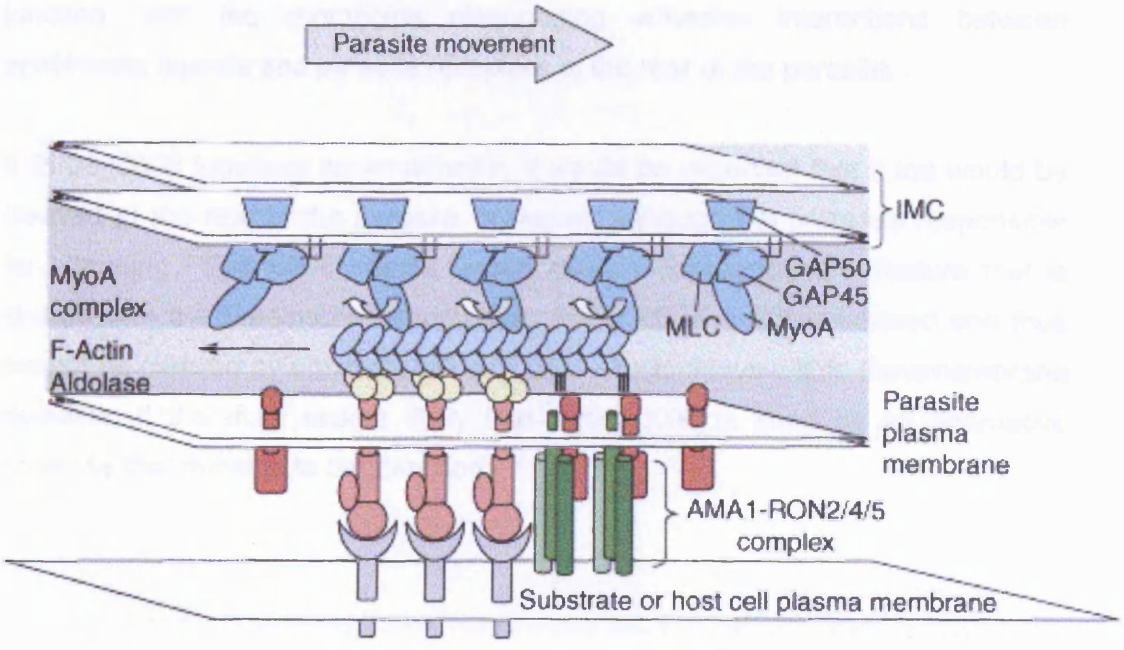
2005), and a comparable complex is believed to exist in *P. falciparum* (Alexander et al., 2006).

The moving junction originates at the tight junction formed between the erythrocyte and the parasite's apical end. Proteins resident in the apical organelles are therefore likely to be components of the moving junction, for example the micronemal protein AMA1 and RON4 which originates from the rhoptry neck. Although PF08_0008 is apically-restricted, there is no evidence from immunoprecipitation experiments to suggest that this protein exists in a complex with any other proteins, for instance, components of the moving junction. However, this cannot be categorically confirmed - interactions with other proteins that are not resistant to the detergent solubilisation conditions employed in immunoprecipitation experiments may indeed exist. If present at the moving junction, the demonstrated binding of PF08_0008 to erythrocytes could in theory serve to exclude a particular host cell molecule from the developing PVM. A model of the apicomplexan moving junction is shown in figure 10.1. The nature of the moving junction - an electron-dense region where the two cells are in intimate contact, makes it extremely difficult to study, and little can be determined from immunofluorescence analyses performed on purified free merozoites, as in the absence of erythrocytes and productive invasion, it is unlikely that they accurately represent what occurs *in vivo*.

***Figure 10.1*: Schematic model of the apicomplexan moving junction.**

Figure adapted from Carruthers and Boothroyd (2007). The parasite's transmembrane adhesins (pink) and their respective ligands in the host cell membrane (grey) are thought to sit behind the moving junction complex, relative to the direction of parasite movement during invasion. These adhesins generate the force for movement via a connection with the molecular motor, which is anchored into the inner membrane complex (IMC) of the cell and consists of aldolase, filamentous actin (F-actin), a myosin light chain (MLC) (MTIP in *P. falciparum*), myosin A (MyoA), GAP45 and GAP50, as discussed in chapter 1.

In *T. gondii*, the moving junction is known to be minimally composed of AMA1 (red), RON2 (medium green), RON4 and RON5 (light and dark green). The arrangement of the AMA1-associated RON proteins shown is hypothetical. The moving junction of *P. falciparum* is also believed to contain the orthologous AMA1 and RON4 proteins, and perhaps further as yet unidentified proteins of which PF08_0008 could be one. It remains to be determined whether any connection between the components of the moving junction and the molecular motor of the parasite exists, or whether any moving junction components have ligands on the host cell.



All known adhesive, microneme-resident proteins that have been described previously are type I integral membrane proteins containing a short cytoplasmic domain. In the proposed molecular mechanism of erythrocyte invasion, parasite proteins that bind erythrocyte surface molecules are indirectly anchored to the actin-myosin motor. Adhesins are presumed to be linked to the motor through an interaction with MTRAP, which has been shown to bind to the F-actin-binding protein aldolase (Baum et al., 2006). It has been shown that in the case of EBA175, the cytoplasmic domain is essential for the molecule to act as an adhesin (Gilberger et al., 2003a), and presumably the motor linkage occurs in this region. In the case of GPI-anchored PF08_0008, there could be no such interaction in the cytoplasm, suggesting that if this protein is linked to the motor complex, it is linked by way of a unique association in the protein ectodomain.

The proteolytic processing and shedding of microneme proteins, such as PF08_0008, is common in *P. falciparum*. However, the parasite-derived enzymes that are responsible for the primary processing and shedding of PF08_0008 are yet to be identified. To date, no parasite proteases responsible for the extensive primary processing of proteins such as MSP1 and AMA1 have been identified. However, three *Plasmodium* enzymes have been identified that are capable of shedding proteins from the surface of the parasite: SUB2, ROM1 and ROM4. SUB2 is believed to be responsible for removing the surface coat at the moving junction, with the rhomboids disengaging adhesive interactions between erythrocyte ligands and parasite receptors at the rear of the parasite.

If PF08_0008 functions as an adhesin, it would be expected that it too would be cleaved at the rear of the parasite. However, although the protease responsible for shedding PF08_0008 seems largely calcium-independent, a feature that is shared with the *Plasmodium* rhomboids, PF08_0008 is GPI-anchored and thus cannot be cleaved by either ROM1 or ROM4 which cleave within transmembrane domains. It therefore seems likely that PF08_0008 is shed by an alternative protease that remains to be identified.

Future work

The data presented implicate PF08_0008 in having a role as an adhesin, but also raise many questions as to how it could function in this capacity.

In order to confirm the erythrocyte-binding function of PF08_0008, a binding partner from erythrocytes needs to be identified. If it is assumed that the erythrocyte receptor is a protein or glycoconjugate, this could be achieved through the use of enzyme-treated erythrocytes that are deficient in particular molecules. For example, treatment of erythrocytes with neuraminidase removes the sialic acid residues from carbohydrate moieties that are required for binding. Another good starting point could be to use overlay assays to estimate the size of any potential erythrocyte-binding partners.

It would also be interesting to determine which region of PF08_0008 is responsible for the erythrocyte-binding activity, especially as PF08_0008 is not predicted to contain any classical protein domains. It has been shown that p42 can be depleted from culture supernatant following the addition of erythrocytes, but since p42 and p37 remain in a complex following primary processing, it is possible that the depletion of p42 is due to its association with p37, which may be the erythrocyte-binding fragment. The erythrocyte-binding region could be determined by using sequential peptides spanning the length of PF08_0008 in erythrocyte-binding assays. The binding of EBA175 to glycophorin A is mediated by N-terminal cysteine rich domains (Sim et al., 1994), and one might therefore speculate that the N-terminal region of PF08_0008 is responsible for its erythrocyte-binding activity, since this is where 7 of the 8 conserved cysteine residues are located.

In order to confirm the sub-cellular location of PF08_0008 and PF08_0008-HA3 in schizonts and free merozoites, it will be necessary to perform immuno-electron microscopy using PF08_0008-specific antibodies, since the small dimensions of the parasite and the close proximity of the rhoptry ducts and micronemes can make it difficult to unambiguously assign parasite proteins to particular sub-cellular compartments by IFA alone. For example, AMA1 was initially thought to be located in the rhoptries (Crewther et al., 1990), but has now been shown to be a micronemal protein by immuno-electron microscopy (Bannister et al., 2003).

It has been shown that the trafficking of the EBL proteins to the micronemes is dependent on a cysteine-rich domain, close to the transmembrane anchor in the protein ectodomain (Treeck et al., 2006). Other micronemal proteins such as AMA1, PTRAMP, MTRAP, SUB2 and now PF08_0008, do not contain such a domain in this location, and this is suggestive of there being multiple mechanisms for microneme targeting. If targeting of PF08_0008 is sequence-dependent, the minimal sequence requirements could be determined using a transgenic approach.

PF08_0008 was initially identified as the protein for study partly because it was detected in both merozoite and sporozoite stages in the proteomic study published in 2002 (Florens et al., 2002). AMA1 was also detected in merozoites and sporozoites, and further investigation showed that it is subject to the same processing and shedding events in both stages (Silvie et al., 2004). It will be interesting to confirm the proteomic data for PF08_0008 and determine if like AMA1, it is expressed, processed and shed in sporozoites as well as merozoites. When searching for a PF08_0008 binding partner on erythrocytes, it may be important to consider the possibility that this protein could perform a host-cell binding function in both erythrocytes and hepatocytes, perhaps binding a generic non-erythrocyte-specific cell-surface molecule.

In order to further investigate whether PF08_0008 exists in the parasite as a part of a complex, as would be expected if it is linked to the motor complex, immunoprecipitations could be carried out using less stringent solubilisation conditions, in order to preserve any interactions that may have been disrupted using the methods utilised in this project.

The transgenic parasites expressing HA-tagged PF08_0008 were originally generated with a view to purifying parasite protein and determining the sites of cleavages responsible for the primary processing and shedding events. However, it is now evident that the protein is GPI-anchored, and that the addition of a triple haemagglutinin tag to the C-terminus has resulted in a soluble rather than membrane-bound PF08_0008. Nevertheless, it appears that the primary processing at least is still occurring as it would in wild-type parasites. The monoclonal antibody 3F10 can therefore hopefully be used in the future to purify the p49 form of PF08_0008-HA3 from parasites, to enable the determination of the primary processing cleavage site, through N-terminal sequencing.

Cleavage site sequence information could also be useful in investigating the identity of the protease responsible for primary processing, perhaps by drawing comparisons with previously elucidated primary processing sites. The importance of this processing for host cell invasion could be also be investigated by attempting the generation of cleavage-resistant mutants, since it is possible that processing 'activates' the protein by revealing a functional, e.g. erythrocyte-binding, domain.

The protease responsible for shedding of PF08_0008 from the surface of the merozoites also remains to be identified, but evidence suggests that it is distinct from ROM1, ROM4, and SUB2. 92 putative proteases were identified in the *P. falciparum* genome, from all major classes including serine, aspartic, cysteine, metallo and threonine (Wu et al., 2003). Repeating the merozoite processing assay with a range of protease inhibitors targeting different functional classes of enzyme could give clues as to the identity of the enzyme responsible for PF08_0008 shedding.

Purified p49 could also be of use in determining if PF08_0008 is exposed to the immune system of infected individuals. Western blots could be performed on purified parasite protein using serum from hyperimmune individuals to determine if this protein is immunogenic. If this is the case, PF08_0008-specific antibodies could in theory be purified from the hyperimmune serum and used in erythrocyte invasion inhibition assays to determine if they are protective.

Given the apparent soluble nature of PF08_0008-HA3, it would be interesting to determine if GPI-anchorage of PF08_0008 is necessary for its redistribution onto the surface of free merozoites, shedding into culture supernatant and erythrocyte-binding ability. EBA175 lacking its membrane anchor has been shown to be shed into culture supernatant in the same manner as wild-type protein, and retains an erythrocyte-binding ability (Reed et al., 2000a), suggesting that membrane anchorage is not essential for adhesins to be discharged from the micronemes, and binding to erythrocyte ligands is still possible, although in the case of EBA175 at least, a membrane anchor is necessary for binding to be functional (Gilberger et al., 2003a).

Attempting to disrupt the endogenous *PF08_0008* gene would reveal if the protein is essential for parasite viability. None of the genes encoding the

DBL-EBP or PfRH proteins is essential for parasite growth, as they can all be successfully disrupted individually (Duraisingh et al., 2003a, Duraisingh et al., 2003b, Gilberger et al., 2003b, Reed et al., 2000a, Stubbs et al., 2005, Triglia et al., 2005). If PF08_0008 also acts as an adhesin, it might be expected that the gene encoding it could be successfully disrupted, although it is important to note that unlike the DBL-EBP and PfRH proteins, PF08_0008 is not encoded by a member of a gene family.

These experiments could all serve to determine if the functioning of PF08_0008 is critical for host cell invasion, and if the protein has any suitability as a future vaccine or chemotherapeutic target.

Acknowledgements

I would like to thank my supervisors Tony Holder and Judith Green, for their advice, support and great ideas throughout the project, especially during the difficult times. For excellent technical help I would like to thank the whole of the Holder lab, in particular Judith Green and Ellen Knuepfer for never tiring of my questions. For training in parasite culture and for providing countless preparations of parasite material I would like to thank Muni Grainger, and Sola Ogun for training and help with the immunisations. I would like to thank Ellen Knuepfer for advising me on all aspects of the parasite work carried out throughout the project, in particular the generation of the transfectants. I thank Irene Ling, Brian Yim Lim, Christine Collins, Mike Blackman, Judith Green and Ellen Knuepfer for antibodies, and Pippa Harris for the transfection vector.

Finally, I would like to thank all of my friends and family for their continual encouragement, without which I would have been lost, in particular my parents, boyfriend Adam, and my friends Nick and Emma.

Bibliography

Aikawa, M., Miller, L. H., Johnson, J. & Rabbege, J. (1978). Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite. *J Cell Biol* **77**, 72-82.

Alexander, D. L., Mital, J., Ward, G. E., Bradley, P. & Boothroyd, J. C. (2005). Identification of the moving junction complex of *Toxoplasma gondii*: a collaboration between distinct secretory organelles. *PLoS Pathog* **1**, e17.

Alexander, D. L., Arastu-Kapur, S., Dubremetz, J. F. & Boothroyd, J. C. (2006). *Plasmodium falciparum* AMA1 binds a rhoptry neck protein homologous to TgRON4, a component of the moving junction in *Toxoplasma gondii*. *Eukaryot Cell* **5**, 1169-73.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403-10.

Amino, R., Thiberge, S., Martin, B., Celli, S., Shorte, S., Frischknecht, F. & Menard, R. (2006). Quantitative imaging of *Plasmodium* transmission from mosquito to mammal. *Nat Med* **12**, 220-4.

Baker, R. P., Wijetilaka, R. & Urban, S. (2006). Two *Plasmodium* Rhomboid Proteases Preferentially Cleave Different Adhesins Implicated in All Invasive Stages of Malaria. *PLoS Pathog* **2**, 922-932.

Bannister, L. H., Butcher, G. A., Dennis, E. D. & Mitchell, G. H. (1975). Structure and invasive behaviour of *Plasmodium knowlesi* merozoites in vitro. *Parasitology* **71**, 483-91.

Bannister, L. H., Mitchell, G. H., Butcher, G. A., Dennis, E. D. & Cohen, S. (1986). Structure and development of the surface coat of erythrocytic merozoites of *Plasmodium knowlesi*. *Cell Tissue Res* **245**, 281-90.

Bannister, L. H. & Dluzewski, A. R. (1990). The ultrastructure of red cell invasion in malaria infections: a review. *Blood Cells* **16**, 257-92; discussion 293-7.

Bannister, L. H., Hopkins, J. M., Fowler, R. E., Krishna, S. & Mitchell, G. H. (2000). A brief illustrated guide to the ultrastructure of *Plasmodium falciparum* asexual blood stages. *Parasitol Today* **16**, 427-33.

Bannister, L. & Mitchell, G. (2003). The ins, outs and roundabouts of malaria. *Trends Parasitol* **19**, 209-13.

Bannister, L. H., Hopkins, J. M., Dluzewski, A. R., Margos, G., Williams, I. T., Blackman, M. J., Kocken, C. H., Thomas, A. W. & Mitchell, G. H. (2003). *Plasmodium falciparum* apical membrane antigen 1 (PfAMA-1) is translocated within micronemes along subpellicular microtubules during merozoite development. *J Cell Sci* **116**, 3825-34.

Baum, J., Richard, D., Healer, J., Rug, M., Krnajski, Z., Gilberger, T. W., Green, J. L., Holder, A. A. & Cowman, A. F. (2006). A conserved molecular motor drives cell invasion and gliding motility across malaria life cycle stages and other apicomplexan parasites. *J Biol Chem* **281**, 5197-208.

Bergman, L. W., Kaiser, K., Fujioka, H., Coppens, I., Daly, T. M., Fox, S., Matuschewski, K., Nussenzweig, V. & Kappe, S. H. (2003). Myosin A tail domain interacting protein (MTIP) localizes to the inner membrane complex of *Plasmodium* sporozoites. *J Cell Sci* **116**, 39-49.

Black, C. G., Wu, T., Wang, L., Hibbs, A. R. & Coppel, R. L. (2001). Merozoite surface protein 8 of *Plasmodium falciparum* contains two epidermal growth factor-like domains. *Mol Biochem Parasitol* **114**, 217-26.

Black, C. G., Wang, L., Wu, T. & Coppel, R. L. (2003). Apical location of a novel EGF-like domain-containing protein of *Plasmodium falciparum*. *Mol Biochem Parasitol* **127**, 59-68.

Blackman, M. J., Heidrich, H. G., Donachie, S., McBride, J. S. & Holder, A. A. (1990). A single fragment of a malaria merozoite surface protein remains on the parasite during red cell invasion and is the target of invasion-inhibiting antibodies. *J Exp Med* **172**, 379-82.

Blackman, M. J., Whittle, H., and Holder, A. A. (1991). Processing of the *Plasmodium falciparum* major merozoite surface protein-1: identification of a 33-kilodalton secondary processing product which is shed prior to erythrocyte invasion. *Mol Biochem Parasitol* **49**, 35-44.

Blackman, M. J., Dennis, E. D., Hirst, E. M., Kocken, C. H., Scott-Finnigan, T. J. & Thomas, A. W. (1996). *Plasmodium knowlesi*: secondary processing of the malaria merozoite surface protein-1. *Exp Parasitol* **83**, 229-39.

Bozdech, Z., Llinas, M., Pulliam, B. L., Wong, E. D., Zhu, J. & DeRisi, J. L. (2003). The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol* **1**, E5.

Brossier, F., Jewett, T. J., Lovett, J. L. & Sibley, L. D. (2003). C-terminal processing of the *Toxoplasma* protein MIC2 is essential for invasion into host cells. *J Biol Chem* **278**, 6229-34.

Brossier, F., Jewett, T. J., Sibley, L. D. & Urban, S. (2005). A spatially localized rhomboid protease cleaves cell surface adhesins essential for invasion by *Toxoplasma*. *Proc Natl Acad Sci U S A* **102**, 4146-51.

Buscaglia, C. A., Coppens, I., Hol, W. G. & Nussenzweig, V. (2003). Sites of interaction between aldolase and thrombospondin-related anonymous protein in *Plasmodium*. *Mol Biol Cell* **14**, 4947-57.

Carruthers, V. B. & Boothroyd, J. C. (2007). Pulling together: an integrated model of *Toxoplasma* invasion. *Curr Opin Microbiol* **10**, 83-9.

Carruthers, V. B. & Sibley, L. D. (1997). Sequential protein secretion from three distinct organelles of *Toxoplasma gondii* accompanies invasion of human fibroblasts. *Eur J Cell Biol* **73**, 114-23.

Carruthers, V. B. & Sibley, L. D. (1999). Mobilization of intracellular calcium stimulates microneme discharge in *Toxoplasma gondii*. *Mol Microbiol* **31**, 421-8.

Carruthers, V. B., Giddings, O. K. & Sibley, L. D. (1999). Secretion of micronemal proteins is associated with *Toxoplasma* invasion of host cells. *Cell Microbiol* **1**, 225-35.

Claros, M. G. & von Heijne, G. (1994). TopPred II: an improved software for membrane protein structure predictions. *Comput Appl Biosci* **10**, 685-6.

Cowman, A. F., and Crabb, B. S. (2006). Invasion of red blood cells by malaria parasites. Cell 124, 755-766.

Crewther, P. E., Culvenor, J. G., Silva, A., Cooper, J. A. & Anders, R. F. (1990). *Plasmodium falciparum*: two antigens of similar size are located in different compartments of the rhoptry. *Exp Parasitol* **70**, 193-206.

Dalley, J. A. & Bulleid, N. J. (2003a). The endoplasmic reticulum (ER) translocon can differentiate between hydrophobic sequences allowing signals for glycosylphosphatidylinositol anchor addition to be fully translocated into the ER lumen. *J Biol Chem* **278**, 51749-57.

Dailey, J. A. & Bulleid, N. J. (2003b). How does the translocon differentiate between hydrophobic sequences that form part of either a GPI (glycosylphosphatidylinositol)-anchor signal or a stop transfer sequence? *Biochem Soc Trans* **31**, 1257-9.

Dame, J. B., Williams, J. L., McCutchan, T. F., Weber, J. L., Wirtz, R. A., Hockmeyer, W. T., Maloy, W. L., Haynes, J. D., Schneider, I., Roberts, D. & et al. (1984). Structure of the gene encoding the immunodominant surface antigen on the sporozoite of the human malaria parasite *Plasmodium falciparum*. *Science* **225**, 593-9.

DePristo, M. A., Zilversmit, M. M. & Hartl, D. L. (2006). On the abundance, amino acid composition, and evolutionary dynamics of low-complexity regions in proteins. *Gene* **378**, 19-30.

Dessens, J. T., Beetsma, A. L., Dimopoulos, G., Wengelnik, K., Crisanti, A., Kafatos, F. C. & Sinden, R. E. (1999). CTRP is essential for mosquito infection by malaria ookinetes. *Embo J* **18**, 6221-7.

Dowse, T. & Soldati, D. (2004). Host cell invasion by the apicomplexans: the significance of microneme protein proteolysis. *Curr Opin Microbiol* **7**, 388-96.

Duffy, M. F., Caragounis, A., Noviyanti, R., Kyriacou, H. M., Choong, E. K., Boysen, K., Healer, J., Rowe, J. A., Molyneux, M. E., Brown, G. V. & Rogerson, S. J. (2006). Transcribed var genes associated with placental malaria in Malawian women. *Infect Immun* **74**, 4875-83.

Duraisingh, M. T., Maier, A. G., Triglia, T. & Cowman, A. F. (2003a). Erythrocyte-binding antigen 175 mediates invasion in *Plasmodium falciparum* utilizing sialic acid-dependent and -independent pathways. *Proc Natl Acad Sci U S A* **100**, 4796-801.

Duraisingh, M. T., Triglia, T., Ralph, S. A., Rayner, J. C., Barnwell, J. W., McFadden, G. I. & Cowman, A. F. (2003b). Phenotypic variation of *Plasmodium falciparum* merozoite proteins directs receptor targeting for invasion of human erythrocytes. *Embo J* **22**, 1047-57.

Dutta, S., Haynes, J. D., Moch, J. K., Barbosa, A. & Lanar, D. E. (2003). Invasion-inhibitory antibodies inhibit proteolytic processing of apical membrane antigen 1 of *Plasmodium falciparum* merozoites. *Proc Natl Acad Sci U S A* **100**, 12295-300.

Dutta, S., Haynes, J. D., Barbosa, A., Ware, L. A., Snavely, J. D., Moch, J. K., Thomas, A. W. & Lanar, D. E. (2005). Mode of action of invasion-inhibitory antibodies directed against apical membrane antigen 1 of *Plasmodium falciparum*. *Infect Immun* **73**, 2116-22.

Eisenhaber, B., Bork, P. & Eisenhaber, F. (1999). Prediction of potential GPI-modification sites in proprotein sequences. *J Mol Biol* **292**, 741-58.

Eisenhaber, B., Bork, P. & Eisenhaber, F. (2001). Post-translational GPI lipid anchor modification of proteins in kingdoms of life: analysis of protein sequence data from complete genomes. *Protein Eng* **14**, 17-25.

Elmendorf, H. G. & Haldar, K. (1993). Identification and localization of ERD2 in the malaria parasite *Plasmodium falciparum*: separation from sites of sphingomyelin synthesis and implications for organization of the Golgi. *Embo J* **12**, 4763-73.

Florens, L., Washburn, M. P., Raine, J. D., Anthony, R. M., Grainger, M., Haynes, J. D., Moch, J. K., Muster, N., Sacci, J. B., Tabb, D. L., Witney, A. A., Wolters, D., Wu, Y., Gardner, M. J., Holder, A. A., Sinden, R. E., Yates, J. R. & Carucci, D. J. (2002). A proteomic view of the *Plasmodium falciparum* life cycle. *Nature* **419**, 520-6.

Foth, B. J., Ralph, S. A., Tonkin, C. J., Struck, N. S., Fraunholz, M., Roos, D. S., Cowman, A. F. & McFadden, G. I. (2003). Dissecting apicoplast targeting in the malaria parasite *Plasmodium falciparum*. *Science* **299**, 705-8.

Frevert, U. (2004). Sneaking in through the back entrance: the biology of malaria liver stages. *Trends Parasitol* **20**, 417-24.

Galinski, M. R., Medina, C. C., Ingravallo, P. & Barnwell, J. W. (1992). A reticulocyte-binding protein complex of *Plasmodium* vivax merozoites. *Cell* **69**, 1213-26.

Galinski, M. R., Xu, M. & Barnwell, J. W. (2000). *Plasmodium* vivax reticulocyte binding protein-2 (PvRBP-2) shares structural features with PvRBP-1 and the *Plasmodium* yoelii 235 kDa rhoptry protein family. *Mol Biochem Parasitol* **108**, 257-62.

Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., Paulsen, I. T., James, K., Eisen, J. A., Rutherford, K., Salzberg, S. L., Craig, A., Kyes, S., Chan, M. S., Nene, V., Shallom, S. J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M. W., Vaidya, A. B., Martin, D. M., Fairlamb, A. H., Fraunholz, M. J., Roos, D. S., Ralph, S. A., McFadden, G. I., Cummings, L. M., Subramanian, G. M., Mungall, C., Venter, J. C., Carucci, D. J., Hoffman, S. L., Newbold, C., Davis, R. W., Fraser, C. M. & Barrell, B. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**, 498-511.

Gerold, P., Schofield, L., Blackman, M. J., Holder, A. A. & Schwarz, R. T. (1996). Structural analysis of the glycosyl-phosphatidylinositol membrane anchor of the merozoite surface proteins-1 and -2 of *Plasmodium falciparum*. *Mol Biochem Parasitol* **75**, 131-43.

Gilberger, T. W., Thompson, J. K., Reed, M. B., Good, R. T. & Cowman, A. F. (2003a). The cytoplasmic domain of the *Plasmodium falciparum* ligand EBA-175 is essential for invasion but not protein trafficking. *J Cell Biol* **162**, 317-27.

Gilberger, T. W., Thompson, J. K., Triglia, T., Good, R. T., Duraisingh, M. T. & Cowman, A. F. (2003b). A novel erythrocyte binding antigen-175 paralogue from *Plasmodium falciparum* defines a new trypsin-resistant receptor on human erythrocytes. *J Biol Chem* **278**, 14480-6.

Gilson, P. R., Nebl, T., Vukcevic, D., Moritz, R. L., Sargeant, T., Speed, T. P., Schofield, L. & Crabb, B. S. (2006). Identification and stoichiometry of glycosylphosphatidylinositol-anchored membrane proteins of the human malaria parasite *Plasmodium falciparum*. *Mol Cell Proteomics* **5**, 1286-99.

Goel, V. K., Li, X., Chen, H., Liu, S. C., Chishti, A. H. & Oh, S. S. (2003). Band 3 is a host receptor binding merozoite surface protein 1 during the *Plasmodium falciparum* invasion of erythrocytes. *Proc Natl Acad Sci U S A* **100**, 5164-9.

Green, J. L., Hinds, L., Grainger, M., Knuepfer, E. & Holder, A. A. (2006). *Plasmodium* thrombospondin related apical merozoite protein (PTRAMP) is shed from the surface of merozoites by PfSUB2 upon invasion of erythrocytes. *Mol Biochem Parasitol* **150**, 114-7.

Green, J. L., Martin, S. R., Fielden, J., Ksagoni, A., Grainger, M., Yim Lim, B. Y., Molloy, J. E. & Holder, A. A. (2006). The MTIP-myosin A complex in blood stage malaria parasites. *J Mol Biol* **355**, 933-41.

Greenwood, B. & Mutabingwa, T. (2002). Malaria in 2002. *Nature* **415**, 670-2.

Guerra, C. A., Snow, R. W. & Hay, S. I. (2006). Mapping the global extent of malaria in 2005. *Trends Parasitol* **22**, 353-8.

Harris, P. K., Yeoh, S., Dluzewski, A. R., O'Donnell, R. A., Withers-Martinez, C., Hackett, F., Bannister, L. H., Mitchell, G. H. & Blackman, M. J. (2005). Molecular identification of a malaria merozoite surface sheddase. *PLoS Pathog* **1**, 241-51.

High, S. & Dobberstein, B. (1992). Mechanisms that determine the transmembrane disposition of proteins. *Curr Opin Cell Biol* **4**, 581-6.

Hofmann, K. & Stoffel, W. (1993). TMbase - A database of membrane spanning proteins segments. *Biological Chemistry Hoppe-Seyler* **374**, 166.

Howell, S. A., Hackett, F., Jongco, A. M., Withers-Martinez, C., Kim, K., Carruthers, V. B. & Blackman, M. J. (2005). Distinct mechanisms govern proteolytic shedding of a key invasion protein in apicomplexan pathogens. *Mol Microbiol* **57**, 1342-56.

Howell, S. A., Well, I., Fleck, S. L., Kettleborough, C., Collins, C. R. & Blackman, M. J. (2003). A single malaria merozoite serine protease mediates shedding of multiple surface proteins by juxtamembrane cleavage. *J Biol Chem* **278**, 23890-8.

Howell, S. A., Withers-Martinez, C., Kocken, C. H., Thomas, A. W. & Blackman, M. J. (2001). Proteolytic processing and primary structure of *Plasmodium falciparum* apical membrane antigen-1. *J Biol Chem* **276**, 31311-20.

Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., De Castro, E., Langendijk-Genevaux, P. S., Pagni, M. & Sigrist, C. J. (2006). The PROSITE database. *Nucleic Acids Res* **34**, D227-30.

Kaslow, D. C., Quakyi, I. A., Syin, C., Raum, M. G., Keister, D. B., Coligan, J. E., McCutchan, T. F. & Miller, L. H. (1988). A vaccine candidate from the sexual stage of human malaria that contains EGF-like domains. *Nature* **333**, 74-6.

Kats, L. M., Black, C. G., Proellocks, N. I. & Coppel, R. L. (2006). *Plasmodium* rhoptries: how things went pear-shaped. *Trends Parasitol* **22**, 269-76.

Kocken, C. H., Jansen, J., Kaan, A. M., Beckers, P. J., Ponnudurai, T., Kaslow, D. C., Konings, R. N. & Schoenmakers, J. G. (1993). Cloning and expression of the gene coding for the transmission blocking target antigen Pfs48/45 of *Plasmodium falciparum*. *Mol Biochem Parasitol* **61**, 59-68.

Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**, 567-80.

Kronegg, J., and Buloz, D. (1999). Detection/predictin of GPI cleavage site (GPI-anchor) in a protein (DGPI). http://129.194.185.165/dgpi/.

Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-32.

LaCount, D. J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J. R., Schoenfeld, L. W., Ota, I., Sahasrabudhe, S., Kurschner, C., et al. (2005). A protein interaction network of the malaria parasite *Plasmodium falciparum*. Nature 438, 103-107.

Lasonder, E., Ishihama, Y., Andersen, J. S., Vermunt, A. M., Pain, A., Sauerwein, R. W., Eling, W. M., Hall, N., Waters, A. P., Stunnenberg, H. G. & Mann, M. (2002). Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. *Nature* **419**, 537-42.

Le Roch, K. G., Zhou, Y., Blair, P. L., Grainger, M., Moch, J. K., Haynes, J. D., De La Vega, P., Holder, A. A., Batalov, S., Carucci, D. J. & Winzeler, E. A. (2003). Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* **301**, 1503-8.

Le Roch, K. G., Johnson, J. R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M., Yan, S. F., Williamson, K. C., Holder, A. A., Carucci, D. J., Yates, J. R., 3rd & Winzeler, E. A. (2004). Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Res* **14**, 2308-18.

Li, X., Chen, H., Oo, T. H., Daly, T. M., Bergman, L. W., Liu, S. C., Chishti, A. H. & Oh, S. S. (2004). A co-ligand complex anchors *Plasmodium falciparum* merozoites to the erythrocyte invasion receptor band 3. *J Biol Chem* **279**, 5765-71.

Ling, I. T., Florens, L., Dluzewski, A. R., Kaneko, O., Grainger, M., Yim Lim, B. Y., Tsuboi, T., Hopkins, J. M., Johnson, J. R., Torii, M., Bannister, L. H., Yates, J. R., 3rd, Holder, A. A. & Mattei, D. (2004). The *Plasmodium falciparum* clag9 gene encodes a rhoptry protein that is transferred to the host erythrocyte upon invasion. *Mol Microbiol* **52**, 107-18.

Llinas, M., Bozdech, Z., Wong, E. D., Adai, A. T. & DeRisi, J. L. (2006). Comparative whole genome transcriptome analysis of three *Plasmodium falciparum* strains. *Nucleic Acids Res* **34**, 1166-73.

Luse, S. A. & Miller, L. H. (1971). *Plasmodium falciparum* malaria. Ultrastructure of parasitized erythrocytes in cardiac vessels. *Am J Trop Med Hyg* **20**, 655-60.

Maier, A. G., Duraisingh, M. T., Reeder, J. C., Patel, S. S., Kazura, J. W., Zimmerman, P. A. & Cowman, A. F. (2003). *Plasmodium falciparum* erythrocyte invasion through glycophorin C and selection for Gerbich negativity in human populations. *Nat Med* **9**, 87-92.

Marshall, V. M., Silva, A., Foley, M., Cranmer, S., Wang, L., McColl, D. J., Kemp, D. J. & Coppel, R. L. (1997). A second merozoite surface protein (MSP-4) of *Plasmodium falciparum* that contains an epidermal growth factor-like domain. *Infect Immun* **65**, 4460-7.

Marshall, V. M., Tieqiao, W. & Coppel, R. L. (1998). Close linkage of three merozoite surface protein genes on chromosome 2 of *Plasmodium falciparum*. *Mol Biochem Parasitol* **94**, 13-25.

Martoglio, B., Hofmann, M. W., Brunner, J. & Dobberstein, B. (1995). The protein-conducting channel in the membrane of the endoplasmic reticulum is open laterally toward the lipid bilayer. *Cell* **81**, 207-14.

Matuschewski, K. (2006). Vaccine development against malaria. *Curr Opin Immunol* **18**, 449-57.

McGuffin, L. J., Bryson, K. & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404-5.

Mendis, K., Sina, B. J., Marchesini, P. & Carter, R. (2001). The neglected burden of *Plasmodium* vivax malaria. *Am J Trop Med Hyg* **64**, 97-106.

Miller, L. H., Baruch, D. I., Marsh, K. & Doumbo, O. K. (2002). The pathogenic basis of malaria. *Nature* **415**, 673-9.

Mitchell, G. H., Thomas, A. W., Margos, G., Dluzewski, A. R. & Bannister, L. H. (2004). Apical membrane antigen 1, a major malaria vaccine candidate, mediates the close attachment of invasive merozoites to host red blood cells. *Infect Immun* **72**, 154-8.

Moran, P. & Caras, I. W. (1994). Requirements for glycosylphosphatidylinositol attachment are similar but not identical in mammalian cells and parasitic protozoa. *J Cell Biol* **125**, 333-43.

Mordue, D. G., Desai, N., Dustin, M. & Sibley, L. D. (1999). Invasion by *Toxoplasma gondii* establishes a moving junction that selectively excludes host cell plasma membrane proteins on the basis of their membrane anchoring. *J Exp Med* **190**, 1783-92.

Naik, R. S., Branch, O. H., Woods, A. S., Vijaykumar, M., Perkins, D. J., Nahlen, B. L., Lal, A. A., Cotter, R. J., Costello, C. E., Ockenhouse, C. F., et al. (2000). Glycosylphosphatidylinositol anchors of *Plasmodium falciparum*: molecular characterization and naturally elicited antibody response that may provide immunity to malaria pathogenesis. J Exp Med 192, 1563-1576.

Narum, D. L. & Thomas, A. W. (1994). Differential localization of full-length and processed forms of PF83/AMA-1 an apical membrane antigen of *Plasmodium falciparum* merozoites. *Mol Biochem Parasitol* **67**, 59-68.

Nebl, T., De Veer, M. J. & Schofield, L. (2005). Stimulation of innate immune responses by malarial glycosylphosphatidylinositol via pattern recognition receptors. *Parasitology* **130 Suppl**, S45-62.

Newbold, C., Craig, A., Kyes, S., Rowe, A., Fernandez-Reyes, D. & Fagan, T. (1999). Cytoadherence, pathogenesis and the infected red cell surface in *Plasmodium falciparum*. *Int J Parasitol* **29**, 927-37.

Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10**, 1-6.

O'Donnell, R. A. & Blackman, M. J. (2005). The role of malaria merozoite proteases in red blood cell invasion. *Curr Opin Microbiol* **8**, 422-7.

O'Donnell, R. A., Hackett, F., Howell, S. A., Treeck, M., Struck, N., Krnajski, Z., Withers-Martinez, C., Gilberger, T. W. & Blackman, M. J. (2006). Intramembrane proteolysis mediates shedding of a key adhesin during erythrocyte invasion by the malaria parasite. *J Cell Biol* **174**, 1023-33.

Ogun, S. A. & Holder, A. A. (1996). A high molecular mass *Plasmodium* yoelii rhoptry protein binds to erythrocytes. *Mol Biochem Parasitol* **76**, 321-4.

O'Keeffe, A. H., Green, J. L., Grainger, M. & Holder, A. A. (2005). A novel Sushi domain-containing protein of *Plasmodium falciparum*. *Mol Biochem Parasitol* **140**, 61-8.

Rayner, J. C., Vargas-Serrato, E., Huber, C. S., Galinski, M. R. & Barnwell, J. W. (2001). A *Plasmodium falciparum* homologue of *Plasmodium* vivax reticulocyte binding protein (PvRBP1) defines a trypsin-resistant erythrocyte invasion pathway. *J Exp Med* **194**, 1571-81.

Reed, M. B., Caruana, S. R., Batchelor, A. H., Thompson, J. K., Crabb, B. S. & Cowman, A. F. (2000a). Targeted disruption of an erythrocyte binding antigen in *Plasmodium falciparum* is associated with a switch toward a sialic acid-independent pathway of invasion. *Proc Natl Acad Sci U S A* **97**, 7509-14.

Reed, M. B., Saliba, K. J., Caruana, S. R., Kirk, K. & Cowman, A. F. (2000b). Pgh1 modulates sensitivity and resistance to multiple antimalarials in *Plasmodium falciparum*. *Nature* **403**, 906-9.

Rieckmann, K. H. (2006). The chequered history of malaria control: are new and better tools the ultimate answer? *Ann Trop Med Parasitol* **100**, 647-62.

Sachs, J. & Malaney, P. (2002). The economic and social burden of malaria. *Nature* **415**, 680-5.

Salanti, A., Staalsoe, T., Lavstsen, T., Jensen, A. T., Sowa, M. P., Arnot, D. E., Hviid, L. & Theander, T. G. (2003). Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering *Plasmodium falciparum* involved in pregnancy-associated malaria. *Mol Microbiol* **49**, 179-91.

Sanders, P. R., Gilson, P. R., Cantin, G. T., Greenbaum, D. C., Nebl, T., Carucci, D. J., McConville, M. J., Schofield, L., Hodder, A. N., Yates, J. R., 3rd & Crabb, B. S. (2005). Distinct protein classes including novel merozoite surface antigens in Raft-like membranes of *Plasmodium falciparum*. *J Biol Chem* **280**, 40169-76.

Schofield, L., Hewitt, M. C., Evans, K., Siomos, M. A. & Seeberger, P. H. (2002). Synthetic GPI as a candidate anti-toxic vaccine in a model of malaria. *Nature* **418**, 785-9.

Schultz, J., Milpetz, F., Bork, P. & Ponting, C. P. (1998). SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* **95**, 5857-64.

Silvie, O., Franetich, J. F., Charrin, S., Mueller, M. S., Siau, A., Bodescot, M., Rubinstein, E., Hannoun, L., Charoenvit, Y., Kocken, C. H., Thomas, A. W., Van Gemert, G. J., Sauerwein, R. W., Blackman, M. J., Anders, R. F., Pluschke, G. & Mazier, D. (2004). A role for apical membrane antigen 1 during invasion of hepatocytes by *Plasmodium falciparum* sporozoites. *J Biol Chem* **279**, 9490-6.

Sim, B. K., Chitnis, C. E., Wasniowska, K., Hadley, T. J. & Miller, L. H. (1994). Receptor and ligand domains for invasion of erythrocytes by *Plasmodium falciparum*. *Science* **264**, 1941-4.

Snow, R. W., Guerra, C. A., Noor, A. M., Myint, H. Y. & Hay, S. I. (2005). The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature* **434**, 214-7.

Spielmann, T. & Beck, H. P. (2000). Analysis of stage-specific transcription in *Plasmodium falciparum* reveals a set of genes exclusively transcribed in ring stage parasites. *Mol Biochem Parasitol* **111**, 453-8.

Spielmann, T., Fergusen, D. J. & Beck, H. P. (2003). etramps, a new *Plasmodium falciparum* gene family coding for developmentally regulated and highly charged membrane proteins located at the parasite-host cell interface. *Mol Biol Cell* **14**, 1529-44.

Spielmann, T., Gardiner, D. L., Beck, H. P., Trenholme, K. R. & Kemp, D. J. (2006). Organization of ETRAMPs and EXP-1 at the parasite-host cell interface of malaria parasites. *Mol Microbiol* **59**, 779-94.

Stubbs, J., Simpson, K. M., Triglia, T., Plouffe, D., Tonkin, C. J., Duraisingh, M. T., Maier, A. G., Winzeler, E. A. & Cowman, A. F. (2005). Molecular mechanism for switching of *P. falciparum* invasion pathways into human erythrocytes. *Science* **309**, 1384-7.

Sturm, A., Amino, R., van de Sand, C., Regen, T., Retzlaff, S., Rennenberg, A., Krueger, A., Pollok, J. M., Menard, R. & Heussler, V. T. (2006). Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science* **313**, 1287-90.

Talman, A. M., Domarle, O., McKenzie, F. E., Ariey, F. & Robert, V. (2004). Gametocytogenesis: the puberty of *Plasmodium falciparum*. *Malar J* **3**, 24.

Thomas, A. W., Bannister, L. H. & Waters, A. P. (1990). Sixty-six kilodalton-related antigens of *Plasmodium knowlesi* are merozoite surface antigens associated with the apical prominence. *Parasite Immunol* **12**, 105-13.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-80.

Tomas, A. M., Margos, G., Dimopoulos, G., van Lin, L. H., de Koning-Ward, T. F., Sinha, R., Lupetti, P., Beetsma, A. L., Rodriguez, M. C., Karras, M., Hager, A., Mendoza, J., Butcher, G. A., Kafatos, F., Janse, C. J., Waters, A. P. & Sinden, R. E. (2001). P25 and P28 proteins of the malaria ookinete surface have multiple and partially redundant functions. *Embo J* **20**, 3975-83.

Topolska, A. E., Lidgett, A., Truman, D., Fujioka, H. & Coppel, R. L. (2004). Characterization of a membrane-associated rhoptry protein of *Plasmodium falciparum*. *J Biol Chem* **279**, 4648-56.

Trape, J. F., Pison, G., Spiegel, A., Enel, C. & Rogier, C. (2002). Combating malaria in Africa. *Trends Parasitol* **18**, 224-30.

Treeck, M., Struck, N. S., Haase, S., Langer, C., Herrmann, S., Healer, J., Cowman, A. F. & Gilberger, T. W. (2006). A conserved region in the EBL-proteins is implicated in microneme targeting of the malaria parasite *Plasmodium falciparum*. *J Biol Chem*.

Triglia, T., Healer, J., Caruana, S. R., Hodder, A. N., Anders, R. F., Crabb, B. S. & Cowman, A. F. (2000). Apical membrane antigen 1 plays a central role in erythrocyte invasion by *Plasmodium* species. *Mol Microbiol* **38**, 706-18.

Triglia, T., Duraisingh, M. T., Good, R. T. & Cowman, A. F. (2005). Reticulocyte-binding protein homologue 1 is required for sialic acid-dependent invasion into human erythrocytes by *Plasmodium falciparum*. *Mol Microbiol* **55**, 162-74.

Trouiller, P. & Olliaro, P. L. (1998). Drug development output from 1975 to 1996: what proportion for tropical diseases? *Int J Infect Dis* **3**, 61-3.

WHO (2000). The African Summit on Roll Back Malaria.

WHO (2003). The World Health Report 2003: shaping the future.

Wu, Y., Wang, X., Liu, X. & Wang, Y. (2003). Data-mining approaches reveal hidden families of proteases in the genome of malaria parasite. *Genome Res* **13**, 601-16.

Zhou, X. W., Blackman, M. J., Howell, S. A. & Carruthers, V. B. (2004). Proteomic analysis of cleavage events reveals a dynamic two-step mechanism for proteolysis of a key parasite adhesive complex. *Mol Cell Proteomics* **3**, 565-76.